HUMAN FRG1 IS AN ACTIN-BUNDLING AND A RNA-ASSOCIATED PROTEIN WITH DISTINCT SUBCELLULAR LOCALIZATIONS

BY

CHIA-YUN SUN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Cell and Developmental Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Associate Professor Michel Bellini, Chair
Assistant Professor Peter L. Jones, Director of Research
Associate Professor Stephanie Ceman
Assistant Professor William M. Brieher
Professor Ann M Nardulli

ABSTRACT

Facioscapulohumeral muscular dystrophy (FSHD) region gene 1 (*FRG1*) is critical for development of the vertebrate musculature and vasculature, however its precise molecular function is unknown. Because of its location 125 kb proximal to the FSHD1A lesion, a deletion in a subtelomeric macrosatellite DNA repeat array, it has been considered a candidate for mediating FSHD pathophysiology. This study investigates FRG1's function to provide insight into FRG1's role in vertebrate muscle development. First, we focus on the subcellular localization of FRG1 and identified FRG1 as a dynamic nuclear, cytoplamic, and sarcomeric protein. During myoblast differentiation, FRG1's subcellular distribution changed dramatically with FRG1 eventually associating with the matured Z-discs. The Z-disc localization is confirmed in mouse myofibers, suggesting FRG1 may have a muscle-specific function involved in sarcomere maintenance or signaling. The nuclear fraction of the endogenous FRG1 is localized in nucleoli, Cajal bodies, and actively transcribed chromatin associated with nascent RNA transcripts, supporting a function in RNA biogenesis. Furthermore, we show that FRG1 interacts specifically with RNA *in vitro*, associates with mRNA *in vivo*, and directly interacts with the TAP mRNA export receptor. FRG1 also exists in a cytoplasmic pool that is dependent on an intact actin cytoskeleton for its localization and we demonstrate FRG1 itself is an actin binding protein. These data provide the first biochemical activities for human FRG1 and indicate that FRG1 dynamically shuttles between the nucleus and cytoplasm and is involved in aspects of RNA biogenesis, potentially including mRNA transport and localization.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION.....1
FSHD PATHOLOGY.....1
FSHD GENETIC DEFECTS.....2
FSHD EPIGENETICS.....3
FSHD DISEASE MODELS.....5
FSHD CANDIDATE GENES.....8
EFFECTS OF ALTERED FRG1 EXPRESSION LEVELS ON DEVELOPMENT.....11
FRG1 IN RNA-BIOGENESIS.....13
STRUCTURE-FUNCTION ANALYSIS OF FRG1.....13
REFERENCES.....15
FIGURES.....21

CHAPTER 2: SUBCELLULAR LOCALIZATIONS OF ENDOGENOUS FRG1.....25
INTRODUCTION.....25
RESULTS.....27
DISCUSSION.....32
MATERIALS AND METHODS.....36
REFERENCES.....42
FIGURES.....44

CHAPTER 3: FRG1 IS A RNA-ASSOCIATED PROTEIN.....51
INTRODUCTION.....51
RESULTS.....53
DISCUSSION.....58
MATERIALS AND METHODS.....60
REFERENCES.....67
FIGURES AND TABLES.....70

CHAPTER 4: FRG1 IS AN ACTIN-BUNDLING PROTEIN.....77
INTRODUCTION.....77
RESULTS.....78
DISCUSSION.....79
MATERIALS AND METHODS.....81
REFERENCES.....84
FIGURES.....85

CHAPTER 5: TRANSLATIONAL IMPACT TOWARD FSHD.....87
REFERENCES.....94
FIGURES.....96

APPENDIX A: PHOSPHORYLATION OF HUMAN FRG1 AFFECTS ITS ACTIN BUNDLING ACTIVITY……97
INTRODUCTION……97
RESULTS……98
DISCUSSION……99
MATERIALS AND METHODS……101
REFERENCES……104
FIGURES AND TABLES……105

CHAPTER 1

INTRODUCTION

FSHD PATHOLOGY

Facioscapulohumeral muscular dystrophy (FSHD) is the most prevalent form of muscular dystrophy in adults (incidence of 1:7,500 – 14,000), and overall the third most common after Duchenne and myotonic muscular dystrophy (2009; Lunt and Harper, 1991; Padberg, 1982). FSHD is inherited in an autosomal dominant manner with approximately 10-30% of affected individuals carrying a *de novo* mutation (Padberg et al., 1995). FSHD is marked by progressive and often asymmetric atrophy in specific muscle groups. As the name implies, FSHD initially affects facial, shoulder, and upper arm muscles, and later extends to pelvic girdle and lower extremities (Olsen et al., 2006; Tawil, 2008). Age of onset is highly variable, ranging from early infancy to late fifties, but most patients become symptomatic in their second decade (Tawil and Van Der Maarel, 2006). There is also a wide variety in progression rate and clinical severity from asymptomatic individuals to patients who eventually require a wheelchair (approximately 20% of affected individuals) (Pandya et al., 2008). In addition, several extramuscular manifestations in FSHD have been described including high frequency hearing loss and retinal telangiectasias in 75% and 60% of patients, respectively (Fitzsimons et al., 1987; Padberg et al., 1995). Although these manifestations are typically asymptomatic, FSHD retinal vasculopathy in severe cases can lead to retinal detachment and blindness, similar to symptoms observed in Coat's disease (Fitzsimons et al., 1987).

FSHD GENETIC DEFECTS

The genetic lesion of FSHD is located on the subtelemeric region of chromosome 4q (4q35.2), where 95% of the patients carry a contraction in an array of 3.3 kb macrosatellite repeats, termed D4Z4 (van Deutekom et al., 1993; Wijmenga et al., 1992). The contraction results in a decrease of repeat number below a threshold of 11, whereas normal individuals carry 11-100 repeats (Figure 1.1). A lower repeat number often correlates to more severe symptoms, although loss of all repeats units does not result in FSHD (Lunt et al., 1995; Lunt et al., 1995). However, there is high degree of interfamilial and intra-familial variability of disease expression with the same size deletion indicating the deletion size is not the only determinant of severity. Interestingly, monosomy of the 4q35 D4Z4 region does not cause FSHD (Tupler et al., 1996), and homozogosity for the FSHD lesion does not result in a more severe phenotype (Tonini et al., 2004) suggesting the lesion results in a gain of function localized to 4q35.

A repeat array that is highly homologous to D4Z4 array was found on chromosome 10q26 (Bakker et al., 1995; Deidda et al., 1995). The D4Z4 array on 10q26 is over 98% homologous to the 4q35 array (Figure 1.1), and the homology extends both proximally (42 kb) and distally to include the telomere (Cacurri et al., 1998; van Geel et al., 2002). This region is also polymorphic with repeat number varying between 1-100 units (van Overveld et al., 2000). Sub-telomeric exchange of the repeat array between 10q26 and 4q35 is observed in 20% of individuals (van Deutekom et al., 1996; van Overveld et al., 2000). However, contraction of D4Z4 repeats array on 10q26 does not cause FSHD suggesting additional *cis* regions specific to chromosome 4q mediate the disease pathophysiology (Bakker et al., 1995; Lemmers et al., 1998).

To further complicate the genetic defect in FSHD, there are two allelic variants of chromosome 4q, 4qA and 4qB (Figure 1.1) (van Geel et al., 2002). The two variants are distinguished by the presence of a 260-bp sequence (pLAM) followed by a 6.2 kb β - satellite repeat in 4qA (Lemmers et al., 2002). Despite the equal frequency of 4qA/B in the population (42% and 58%, respectively), FSHD is exclusively associated with the 4qA type (Lemmers et al., 2002; Lemmers et al., 2004). Using a combination of polymorphic markers, including a single nucleotide polymorphism (SNP) within the most proximal copy of D4Z4 unit, a simple sequence length polymorphism (SSLP) that's 3.5 kb proximal to the repeat array (Figure 1.2), and the A/B variation, subtelomeric region of chromosome 4q can be subdivided into at least 9 haplotypes (Lemmers et al., 2007). There are three haplotypes within 4qA, of which only 4qA161 is found to be associated with FSHD. This allele is nonpathogenic since 4qA161 haplotype was found in 86% of normal individuals with 4qA allele, however, it is FSHD permissive since it is required in combination with the D4Z4 deletion for disease progression (Lemmers et al., 2007).

FSHD EPIGENETICS

Since the FSHD genetic defect is not a classical mutation within a known protein-coding gene, it's been postulated that the contraction results in mis-regulation of 4q35 genes through an epigenetic mechanism. To understand the chromatin structure of 4q35 D4Z4 array, the field has studied this region in respects of DNA methylation, histone modifications, and nuclear organization.

D4Z4 repeat units have an unusually high GC content (73%) with characteristics of CpG islands and containing two repetitive sequences, LSau and hhspm3, which are

predominantly found in heterochromatic regions (Hewitt et al., 1994; Lee et al., 1995). Generally, DNA methylation on CpG islands correlates to increased chromatin condensation and gene silencing. The D4Z4 region in individuals with FSHD is marked by DNA hypomethylation compared to healthy individuals. Furthermore, individuals with phenotypic FSHD (FSHD2) without a contracted D4Z4 (~5% of FSHD patients) also have DNA hypomethylation of this D4Z4 region (Tsien et al., 2001; van Overveld et al., 2003). These data support that changes in the epigenetic landscape of 4q35 may lead to FSHD pathophysiology.

The chromatic nature of 4q35 was also characterized by chromatin immunoprecipitation assay (ChIP) using an antibody specific for acetylated histone H4, a marker for transcriptionally active chromatin. The H4 acetylation levels of the region adjacent to 4q35 D4Z4 arrays were like those found in unexpressed euchromatin rather than the levels indicative of constitutive heterochromatin found in normal and FSHD lymphoid cells (Jiang et al., 2003). Fluorescence *in situ* hybridization (FISH) analysis of 4q35 using cultured myoblasts from both FSHD patients and healthy individuals showed that 4q35 region does not co-localize with heterochromatic markers, nor with histone H4 lysine 8 (H4K8) acetylation, a marker for highly expressed euchromatin, (Jiang et al., 2003). The 4q35 FISH results confirmed that 4q35 region resembles unexpressed euchromatin rather than constitutive heterochromatin meaning the region may be more susceptible to epigenetic changes affecting localized gene expression.

Mammalian chromosomes occupy distinct territories in the nucleus; and many nuclear functions, such as transcription, RNA-processing, and replication, take place in well-defined nuclear compartments (Francastel et al., 2000; Manuelidis, 1984). Unlike

other chromosome ends studied, the FSHD region near the chromosome 4q telomere is closely associated with the nuclear periphery in several cell types (Masny et al., 2004). The nuclear envelope and associated lamina are the nuclear domain that play a role in gene expression, chromatin organization, and differentiation (Marshall, 2003). Interestingly, several neuromuscular disorders result from deficiencies in components of the nuclear envelope, such as Emery-Dreifuss muscular dystrophy (EDMD) (Burke and Stewart, 2002; Ostlund and Worman, 2003). This finding suggests that FSHD may arise through aberrant interactions of distal 4q chromatin with the nuclear envelope.

FSHD DISEASE MODELS

Several models have emerged to explain how the contraction in D4Z4 array leads to FSHD pathogenesis. These include the loss of silencing of a transcript within each D4Z4 called *DUX4*, loss of silencing of proximal 4q35 genes similar to the position effect variegation (PEV), altered nuclear organization model, and chromosome looping model.

D4Z4 gene expression model

A putative double homeodomain gene, *DUX4*, was identified within each D4Z4 repeat unit. It was proposed that the partial deletion of the D4Z4 repeat array resulted in destabilization of the D4Z4 heterochromatin and the inappropriate upregulation of DUX4 (Figure 1.3A) (Gabriels et al., 1999; Hewitt et al., 1994). The functionality of the *DUX4* gene has been questioned due to the lack of introns and polyadenylation signals, and absence of transcription *in vivo* (Alexiadis et al., 2007; Gabriels et al., 1999; Hewitt et al., 1994; Osborne et al., 2007; Winokur et al., 2003). However, recent studies showed that

the open reading frame (ORF) of *DUX4* is conserved in several mammalian species suggesting a coding role, and expression of mouse *Dux4* genes was detected (Clapp et al., 2007). Further, expression of two different *DUX4* transcripts in cells transfected with D4Z4 elements and in FSHD myoblasts were reported (Kowaljow et al., 2007). This transcript contains two introns and is transcribed from the most distal D4Z4 repeat unit (Dixit et al., 2007). Interestingly, the pLAM sequence immediately distal to the D4Z4 array, which is found in the 4qA haplotype mentioned above, may provide the polyadenylation signal for the transcript. This distal *DUX4* transcript can only be observed in FSHD myotubes but not in control myotubes.

Cis-spreading model

Another model predicts an indirect role for the D4Z4 contraction in FSHD pathogenesis, in which D4Z4 contraction may alter the chromatin structure of the array, like D4Z4 hypomethylation, and lead to loss of silencing of candidate gene *in cis* (Figure 1.3B). A D4Z4 binding element (DBE) within the D4Z4 repeat unit was identified, which recruits a repressor complex consists of YY1, a transcription factor that can act as transcription activator or repressor; HMGB2, an chromatin architectural protein; and RNA-binding protein nucleolin, which is involved in transcriptional control of ribosomal RNA and in ribogenesis (Gabellini et al., 2002). This repressor complex was shown to bind to D4Z4 *in vitro* and *in vivo*, and mediate transcriptional repression of 4q35 genes; 4q35 genes, *FRG2, FRG1,* and *ANT1*, were found transcriptionally upregulated in a distance-dependent manner in FSHD muscle samples. However, further expression analysis could not confirm the unregulation of these 4qter genes in FSHD samples (Jiang

et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; Winokur et al., 2003). Further, the normal 4q35 region is not constitutive heterochromatin as previously described, and this model failed to explain why FSHD is only linked to the 4qA phenotype while contraction on 4qB is non-pathogenic.

Altered chromosome looping model

A nuclear matrix attachment site (S/MAR) was identified and located in the proximal vicinity of the D4Z4 repeat array, which is important for the organization of DNA into loop domains as part of a higher order chromatin structure (Razin et al., 2007). This S/MAR site was found to be prominent in normal human myoblasts, however, much weaker in muscle cells derived from FSHD patients, suggesting that the D4Z4 repeat array and upstream genes reside in two loops in normal human myoblasts but in only one loop in FSHD (Figure 1.3C) (Petrov et al., 2006). Further, a strong transcriptional enhancer resides in the 5' end of the D4Z4 repeat unit was identified, therefore, altering chromosome looping in this region may results in the upregulation of neighboring genes (Petrov et al., 2008).

Nuclear organization model

FSHD region, the subtelomeric region of chromosome 4q, is closely associated with the nuclear periphery across several cell types as previously described (Masny et al., 2004). Importantly, this localization seems to be dependent on the presence of lamin A/C, a major component of the nuclear lamina underlying the nuclear envelope, since this peripheral localization of 4qter is lost in cells deficient of lamin A/C (Masny et al., 2004).

No disturbed localization was observed for the disease allele in myoblasts derived from FSHD patients compared to controls (Masny et al., 2004). An independent study has also demonstrated the intrinsic properties of 4qter are necessary and sufficient to localize this subtelomeric domain in the nuclear periphery (Tam et al., 2004). In this nuclear localization model, FSHD may arise through aberrant interactions of distal 4q chromatin with transcription factors or chromatin modifiers at the nuclear envelope, leading to change of transcriptional regulation of proximal genes (Figure 1.3D).

FSHD CANDIDATE GENES

Although the exact disease mechanism of FSHD is unclear, FSHD D4Z4 deletion likely results in the epigenetic misregulation of one or more 4q35 localized genes (de Greef et al., 2008; van der Maarel and Frants, 2005). There are multiple FSHD candidate genes in proximity to the 4q35 D4Z4 array, including FSHD region gene 1 (*FRG1*), FSHD region gene 2 (*FRG2*), adenine nucleotide translocator 1 (*ANT1*), *DUX4c*, and one putative gene within the array, *DUX4* (Figure 1.4).

FRG2 and ANT1

FRG2, located 37 kb proximal to the D4Z4 array, has no known function, has never been shown to produce a protein *in vivo*, and is not conserved beyond humans (Klooster et al., 2009). However, FRG2 is transcriptionally upregulated in differentiating primary myoblast cultures derived from FSHD patients (Rijkers et al., 2004). Interestingly, it was shown that the upregulation of FRG2 mainly originated from chromosome 10q26 rather than 4q35. Like other candidate genes, FRG2 misregulation in

FSHD is controversial (Jiang et al., 2003; Osborne et al., 2007; Winokur et al., 2003), and expression analysis failed to detect FRG2 transcripts in biopsy samples from FSHD patient or normal controls (Klooster et al., 2009). Transgenic mice over-expressing FRG2 did not develop a muscular dystrophy phenotype (Gabellini et al., 2006). Furthermore, patients with a deletion that extends into the proximal non-repeated region including FRG2 are clinically indistinguishable from FSHD patients indicating FRG2 gain of expression is not involved in FSHD (Deak et al., 2007; Lemmers et al., 2003).

The *ANT1* gene encodes a conserved mitochondrial protein that facilitates export of ATP over the mitochondrial membrane. ANT1 levels were shown to be increased in both unaffected and affected FSHD muscles compared to controls. Increased expression level of ANT1 may sensitize cells to oxidative stress and apoptosis (Laoudj-Chenivesse et al., 2005). Misregulation of *ANT1* in FSHD is also controversial and transgenic mice overexpressing ANT1 did not have an FSHD-like phenotype (Gabellini et al., 2006; Gabellini et al., 2002; Tupler et al., 1999).

DUX4 and DUX4c

A putative gene was identified within each D4Z4 unit designated *DUX4* (double homeobox gene 4). *DUX4* expression was detected in FSHD myoblasts but not in normal myoblasts; and expression of mouse *Dux4* gene was reported (Clapp et al., 2007; Kowaljow et al., 2007). DUX4 protein expression may induce apoptosis and caspase 3/7 activity, and alters emerin distribution at the nuclear envelope (Kowaljow et al., 2007). Another study reported DUX4 could activate transient expression of a luciferase reporter gene fused to the Pitx1 promoter as well as the endogenous Pitx1 gene (paired-like

homeodomain transcription factor 1) in transfected C2C12 cells (Dixit et al., 2007). PITX1 is a member of a paired family of homeoproteins and is involved in specification of hind limb identity. PITX1 was found to be specifically upregulated in patients with FSHD compared to controls (Dixit et al., 2007). However, recent expression analysis could not confirm the overexpression of PITX1 (Klooster et al., 2009; Masny et al., 2010). Further, both DUX4 and PITX1 appear to be extremely toxic for cells; expression of even very low levels of either DUX4 or pitx1 in *X. laevis* led to massive cellular loss and severely abnormal development, and these abnormalities were not muscle specific (Wuebbles et al., 2010). More studies are required in order to determine if DUX4 is involved in FSHD pathogenesis.

An inverted copy of *DUX4* was found approximately 40 kb proximal to D4Z4 array, termed *DUX4c* (Double homeobox 4 centromeric) (Wright et al., 1993). DUX4c protein expression was detected in primary myoblasts, and a 2-10 fold upregulation was observed in FSHD myoblasts and muscle biopsies compared to control (Ansseau et al., 2009). *DUX4c* overexpression was shown to induce MYF5 expression, a myogenic regulator, which is known to inhibit myoblast differentiation (Ansseau et al., 2009). However, extended proximal deletion of the D4Z4 array including *DUX4c* was reported in some FSHD patients makes *DUX4c* a less likely candidate gene for FSHD.

FRG1

FRG1 was one of the earliest candidate genes for FSHD since it was the first gene identified in the 4q35 region (van Deutekom et al., 1996). *FRG1* is located 125 kb centromeric to D4Z4 array, and is very highly conserved from vertebrates to invertebrates

(human FRG1 shares 97%, 81%, and 46% homology with mouse, *Xenopus*, and *C. elegans*, respectively; see Figure 1.5) suggesting important biological functions that were conserved throughout evolution (Grewal et al., 1998). Like other FSHD candidate genes, multiple expression profile studies failed to show consistent results that FRG1 is mis-regulated in FSHD samples, with expression changes varying from 25-fold increased, unchanged, and 5-fold decreased compared with controls (Gabellini et al., 2002; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; van Deutekom et al., 1996). However, unlike FRG2 and ANT1, altering FRG1 expression levels have led to an FSHD-like phenotype in multiple animal models including mouse, *Xenopus*, and *C. elegans* (Gabellini et al., 2006; Hanel et al., 2009; Liu et al., 2010; Wuebbles et al., 2009). Significantly, among the FSHD candidate genes, *FRG1* is the only one that has been shown to be required for normal development of both the vasculature and the musculature, the two affected tissues in FSHD, and its overexpression leads to both muscular and vascular FSHD-like phenotypes (Hanel et al., 2009; Wuebbles et al., 2009) as discussed below. Thus, *FRG1* is the most viable FSHD candidate gene.

EFFECTS OF ALTERED FRG1 EXPRESSION LEVELS ON DEVELOPMENT

Protein and mRNA expression profile studies have failed to determine conclusively if FRG1 is mis-regulated in FSHD. How, or if, FRG1 misexpression could mediate FSHD disease pathogenesis is unknown due to the lack of understanding of FRG1's molecular function. However, altering expression levels of FRG1 in animal models have led to disrupted musculature and vasculature, the two prominent features of FSHD.

FRG1 is critical for muscle development

Transgenic mice over-expressing FRG1 (with a 25-40 fold increase) in skeletal muscle results in an FSHD-like phenotype: smaller muscle fibers, reduction in body weight, and muscle atrophy (Gabellini et al., 2006). Altering expression levels of frg1 in developing frogs (*Xenopus laevis*) revealed frg1 is critical for muscle development. More specifically, decreasing frg1 protein levels by antisense morpholino microinjection disrupted myotome organization and inhibited myotome growth; increasing frg1 protein levels by mRNA injection led to abnormal epaxial and hypaxial muscle formation (Hanel et al., 2009). In addition, overexpressing FRG-1 in *C. elegans* specifically from the *frg-1* promoter disrupts the adult ventral muscle structure and organization (Liu et al., 2010). All results support that maintaining normal levels of FRG1 is critical for muscle development.

FRG1 is crucial for developing vasculature

In addition to muscular symptoms, 50-75% of FSHD patients also exhibit retinal vasculopathy (Fitzsimons et al., 1987). Characterizing endogenous frg1 expression in *X. laevis* using an antibody specific for frg1 revealed that frg1 is expressed in both the developing and adult vasculature. Decreasing frg1 levels results in decreased angiogenesis, while overexpressing FRG1 lead to pro-angiogenic phenotypes of increased blood vessel branching and dilation (Wuebbles et al., 2009). These results showed that mis-regulation of FRG1 leads to disrupted vascular phenotype, which is the first link between a 4q35 candidate gene and the vasculature phenotype in FSHD pathology.

FRG1 IN RNA-BIOGENESIS

Very little is known about the FRG1 protein and no biological activity has been described for FRG1. Domain analysis of FRG1 revealed two nuclear localization signals (NLS) residing in the N-terminus, one bipartite NLS domain in the C-terminus, and a fascin-like domain, found in actin-bundling and cross-linking proteins in the middle (Kureishy et al., 2002; van Koningsbruggen et al., 2004). Several lines of evidence suggest FRG1 is involved in RNA biogenesis. Epitope-tagged FRG1 overexpressed in vertebrate cell culture appears completely nuclear, and localized to the nucleoli, Cajal bodies, and nuclear speckles, sites where RNA-biogenesis is taking place (van Koningsbruggen et al., 2004). FRG1 has been found to associate with RNA processing components and large-scale proteomic studies revealed that FRG1 is associated with human spliceosome complex (Rappsilber et al., 2002; van Koningsbruggen et al., 2007). Furthermore, mis-spliced mRNA transcripts were found in transgenic mice overexpressing Frg1 (Gabellini et al., 2002). However, there is no direct evidence of FRG1's specific role in RNA biogenesis, and the subcellular localization of the endogenous human FRG1 has not been characterized.

STUCTURE-FUNCTION ANALYSIS OF FRG1

Whether or not FRG1 is involved in FSHD pathogenesis is not clear due in part to the lack of understanding of FRG1's molecular function. Despite the proposed role of FRG1 in RNA-biogenesis, the precise function of FRG1 remains unknown. Here, we want to focus on the function of human FRG1. First, subcellular localization of endogenous human FRG1 was investigated in multiple cell types. To determine if FRG1

plays a role in muscle differentiation, the subcellular localization of FRG1 was monitored during myogenesis using cultured human skeletal muscle myoblasts. In addition, the nuclear localization of endogenous FRG1 was characterized in order to further investigate FRG1's role in RNA biogenesis. The potential direct association of FRG1 with RNA was investigated *in vivo* and *in vitro* using RNA-immunoprecipitation (RNA-IP) and RNA binding assay, respectively. Finally, the fascin domain in FRG1 suggests FRG1 may bundle actin or associate with the actin cytoskeleton. The human FRG1 actin-bundling activities were characterized and the localization of endogeneous FRG1 was investigated in cells with disrupted actin cytoskeleton.

REFERENCES

(2009). "prevalence of rare diseases: bibliographic data." from http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_alphabetical_list.pdf.

Alexiadis, V., M. E. Ballestas, et al. 2007. RNAPol-ChIP analysis of transcription from FSHD-linked tandem repeats and satellite DNA. Biochim Biophys Acta, **1769**: 29-40.

Ansseau, E., D. Laoudj-Chenivesse, et al. 2009. DUX4c is up-regulated in FSHD. It induces the MYF5 protein and human myoblast proliferation. PLoS One, **4**: e7482.

Bakker, E., C. Wijmenga, et al. 1995. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. Muscle Nerve, **2**: S39-44.

Burke, B. and C. L. Stewart. 2002. Life at the edge: the nuclear envelope and human disease. Nat Rev Mol Cell Biol, **3**: 575-85.

Cacurri, S., N. Piazzo, et al. 1998. Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric KpnI repeat units implicated in facioscapulohumeral muscular dystrophy. Am J Hum Genet, **63**: 181-90.

Clapp, J., L. M. Mitchell, et al. 2007. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. Am J Hum Genet, **81**: 264-79.

de Greef, J. C., R. R. Frants, et al. 2008. Epigenetic mechanisms of facioscapulohumeral muscular dystrophy. Mutat Res, **647**: 94-102.

Deak, K. L., R. J. Lemmers, et al. 2007. Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. Neurology, **68**: 578-82.

Deidda, G., S. Cacurri, et al. 1995. Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. Eur J Hum Genet, **3**: 155-67.

Dixit, M., E. Ansseau, et al. 2007. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. Proc Natl Acad Sci U S A, **104**: 18157-62.

Fitzsimons, R. B., E. B. Gurwin, et al. 1987. Retinal vascular abnormalities in facioscapulohumeral muscular dystrophy. A general association with genetic and therapeutic implications. Brain, **110 (Pt 3)**: 631-48.

Francastel, C., D. Schubeler, et al. 2000. Nuclear compartmentalization and gene activity. Nat Rev Mol Cell Biol, **1**: 137-43.

Gabellini, D., G. D'Antona, et al. 2006. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. Nature, **439**: 973-7.

Gabellini, D., M. R. Green, et al. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell, **110**: 339-48.

Gabriels, J., M. C. Beckers, et al. 1999. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. Gene, **236**: 25-32.

Grewal, P. K., L. C. Todd, et al. 1998. FRG1, a gene in the FSH muscular dystrophy region on human chromosome 4q35, is highly conserved in vertebrates and invertebrates. Gene, **216**: 13-9.

Hanel, M. L., R. D. Wuebbles, et al. 2009. Muscular dystrophy candidate gene FRG1 is critical for muscle development. Dev Dyn, **238**: 1502-12.

Hewitt, J. E., R. Lyle, et al. 1994. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. Hum Mol Genet, **3**: 1287-95.

Jiang, G., F. Yang, et al. 2003. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet, **12**: 2909-21.

Klooster, R., K. Straasheijm, et al. 2009. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. Eur J Hum Genet, **17**: 1615-24.

Kowaljow, V., A. Marcowycz, et al. 2007. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. Neuromuscul Disord, **17**: 611-23.

Kureishy, N., V. Sapountzi, et al. 2002. Fascins, and their roles in cell structure and function. Bioessays, **24**: 350-61.

Laoudj-Chenivesse, D., G. Carnac, et al. 2005. Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. J Mol Med, **83**: 216-24.

Lee, J. H., K. Goto, et al. 1995. Characterization of a tandemly repeated 3.3-kb KpnI unit in the facioscapulohumeral muscular dystrophy (FSHD) gene region on chromosome 4q35. Muscle Nerve, **2**: S6-13.

Lemmers, R. J., P. de Kievit, et al. 2002. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. Nat Genet, **32**: 235-6.

Lemmers, R. J., M. Osborn, et al. 2003. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. Neurology, **61**: 178-83.

Lemmers, R. J., S. M. van der Maarel, et al. 1998. Inter- and intrachromosomal subtelomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. Hum Mol Genet, **7**: 1207-14.

Lemmers, R. J., M. Wohlgemuth, et al. 2004. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. Am J Hum Genet, **75**: 1124-30.

Lemmers, R. J., M. Wohlgemuth, et al. 2007. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. Am J Hum Genet, **81**: 884-94.

Liu, Q., T. I. Jones, et al. 2010. Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. J Cell Sci, **123**: 1116-23.

Lunt, P. W. and P. S. Harper. 1991. Genetic counselling in facioscapulohumeral muscular dystrophy. J Med Genet, **28**: 655-64.

Lunt, P. W., P. E. Jardine, et al. 1995. Phenotypic-genotypic correlation will assist genetic counseling in 4q35-facioscapulohumeral muscular dystrophy. Muscle Nerve, **2**: S103-9.

Lunt, P. W., P. E. Jardine, et al. 1995. Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). Hum Mol Genet, **4**: 951-8.

Manuelidis, L. 1984. Different central nervous system cell types display distinct and nonrandom arrangements of satellite DNA sequences. Proc Natl Acad Sci U S A, **81**: 3123-7.

Marshall, W. F. 2003. Gene expression and nuclear architecture during development and differentiation. Mech Dev, **120**: 1217-30.

Masny, P. S., U. Bengtsson, et al. 2004. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? Hum Mol Genet, **13**: 1857-71.

Masny, P. S., O. Y. Chan, et al. 2010. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. Eur J Hum Genet, **18**: 448-56.

Olsen, D. B., P. Gideon, et al. 2006. Leg muscle involvement in facioscapulohumeral muscular dystrophy assessed by MRI. J Neurol, **253**: 1437-41.

Osborne, R. J., S. Welle, et al. 2007. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology, **68**: 569-77.

Ostlund, C. and H. J. Worman. 2003. Nuclear envelope proteins and neuromuscular diseases. Muscle Nerve, **27**: 393-406.

Padberg, G. W. 1982. Facioscapulohumeral Disease.

Padberg, G. W., O. F. Brouwer, et al. 1995. On the significance of retinal vascular disease and hearing loss in facioscapulohumeral muscular dystrophy. Muscle Nerve, **2**: S73-80.

Padberg, G. W., R. R. Frants, et al. 1995. Facioscapulohumeral muscular dystrophy in the Dutch population. Muscle Nerve, **2**: S81-4.

Pandya, S., W. M. King, et al. 2008. Facioscapulohumeral dystrophy. Phys Ther, **88**: 105-13.

Petrov, A., J. Allinne, et al. 2008. A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy. Genome Res, **18**: 39-45.

Petrov, A., I. Pirozhkova, et al. 2006. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. Proc Natl Acad Sci U S A, **103**: 6982-7.

Rappsilber, J., U. Ryder, et al. 2002. Large-scale proteomic analysis of the human spliceosome. Genome Res, **12**: 1231-45.

Razin, S. V., O. V. Iarovaia, et al. 2007. Chromatin domains and regulation of transcription. J Mol Biol, **369**: 597-607.

Rijkers, T., G. Deidda, et al. 2004. FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. J Med Genet, **41**: 826-36.

Tam, R., K. P. Smith, et al. 2004. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres. J Cell Biol, **167**: 269-79.

Tawil, R. 2008. Facioscapulohumeral muscular dystrophy. Neurotherapeutics, **5**: 601-6.

Tawil, R. and S. M. Van Der Maarel. 2006. Facioscapulohumeral muscular dystrophy. Muscle Nerve, **34**: 1-15.

Tonini, M. M., R. C. Pavanello, et al. 2004. Homozygosity for autosomal dominant facioscapulohumeral muscular dystrophy (FSHD) does not result in a more severe phenotype. J Med Genet, **41**: e17.

Tsien, F., B. Sun, et al. 2001. Methylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cell cultures and tissues. Mol Genet Metab, **74**: 322-31.

Tupler, R., A. Berardinelli, et al. 1996. Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. J Med Genet, **33**: 366-70.

Tupler, R., G. Perini, et al. 1999. Profound misregulation of muscle-specific gene expression in facioscapulohumeral muscular dystrophy. Proc Natl Acad Sci U S A, **96**: 12650-4.

van der Maarel, S. M. and R. R. Frants. 2005. The D4Z4 repeat-mediated pathogenesis of facioscapulohumeral muscular dystrophy. Am J Hum Genet, **76**: 375-86.

van Deutekom, J. C., E. Bakker, et al. 1996. Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. Hum Mol Genet, **5**: 1997-2003.

van Deutekom, J. C., R. J. Lemmers, et al. 1996. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. Hum Mol Genet, **5**: 581-90.

van Deutekom, J. C., C. Wijmenga, et al. 1993. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. Hum Mol Genet, **2**: 2037-42.

van Geel, M., M. C. Dickson, et al. 2002. Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. Genomics, **79**: 210-7.

van Koningsbruggen, S., R. W. Dirks, et al. 2004. FRG1P is localised in the nucleolus, Cajal bodies, and speckles. J Med Genet, **41**: e46.

van Koningsbruggen, S., K. R. Straasheijm, et al. 2007. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. Chromosoma, **116**: 53-64.

van Overveld, P. G., R. J. Lemmers, et al. 2000. Interchromosomal repeat array interactions between chromosomes 4 and 10: a model for subtelomeric plasticity. Hum Mol Genet, **9**: 2879-84.

van Overveld, P. G., R. J. Lemmers, et al. 2003. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nat Genet, **35**: 315-7.

Wijmenga, C., J. E. Hewitt, et al. 1992. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. Nat Genet, **2**: 26-30.

Winokur, S. T., Y. W. Chen, et al. 2003. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. Hum Mol Genet, **12**: 2895-907.

Wright, T. J., C. Wijmenga, et al. 1993. Fine mapping of the FSHD gene region orientates the rearranged fragment detected by the probe p13E-11. Hum Mol Genet, **2**: 1673-8.

Wuebbles, R. D., M. L. Hanel, et al. 2009. FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. Dis Model Mech, **2**: 267-74.

Wuebbles, R. D., S. W. Long, et al. 2010. Testing the effects of FSHD candidate gene expression in vertebrate muscle development. Int J Clin Exp Pathol, **3**: 386-400.

FIGURES

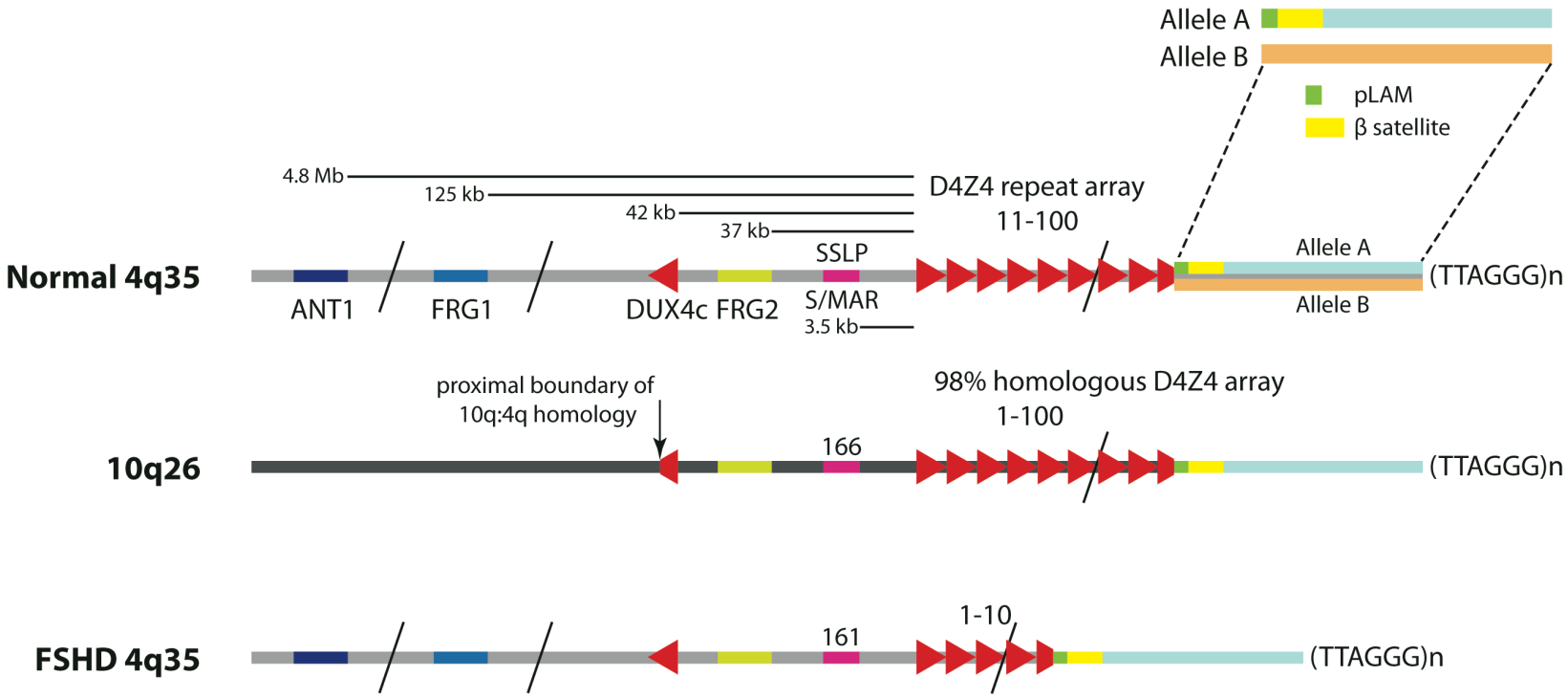

Figure 1.1: Genetic defect of FSHD is a contraction of repeat element D4Z4 on chromosome 4q. Schematic diagram shows the normal 4q35 region with two allelic variants in the 4q subtelomere. A highly homologous region was found on 10q26; however, FSHD is specifically linked to contractions on 4q35 with subtelomeric allele A and SSLP 161.

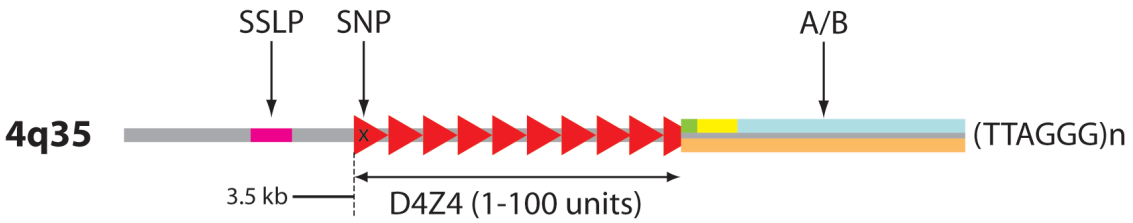

Figure 1.2: Subtelomeric region of chromosome 4q can be subdivided into several haplotypes. Simple sequence length polymorphism (SSLP), SNP within the most proximal D4Z4 unit, and A/B variation were used as polymorphism markers to distinguish the allelic haplotypes within chromosome 4q.

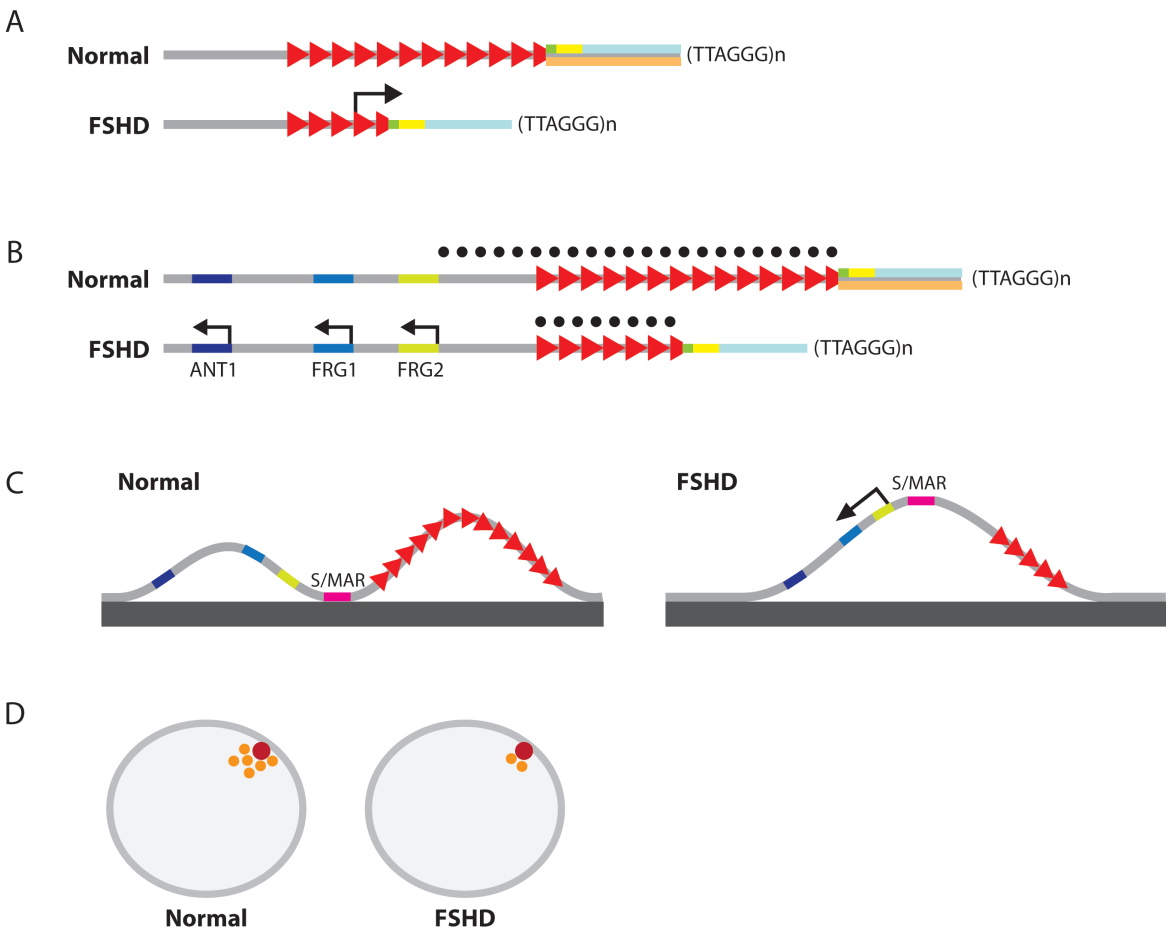

Figure 1.3: Disease models proposed for FSHD. Several disease models have been proposed to explain how the contraction in D4Z4 leads to FSHD pathogenesis. These include the D4Z4 gene expression model (A), *Cis*-spreading model (B), altered chromosome looping model (C), and nuclear organization model (D).

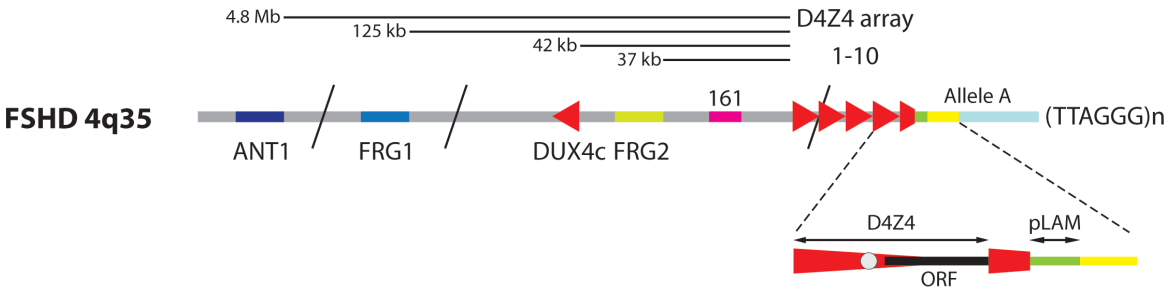

Figure 1.4: Multiple candidate genes proposed for FSHD. There are multiple FSHD candidate genes in proximity to or within the 4q35 D4Z4 array. The putative ORF within D4Z4 is indicated, and D4Z4 binding element (DBE) is represented by the gray circle.

```
                                                    NLS                 Lipocalin
H. sapiens    MAEYSYVKSTKLVLKGTKTKSKKKKSKDKKRKREEDEETQLDIV--GIWWTVTNFGEISG--TIAIEM
M. musculus   MAEYSYVKSTKLVLKGTKAKSKKKKSKDKKRKREEDEETQLDIV--GIWWTVSNFGEISG--TIAIEM
X. laevis     MAEYSKVKSTKLMLKGMKNKSK--KNKDKKRKREESDEDKLDIA--GNWWSVKNFGEISG--TVAIEM
C. elegans MPGADYSAVKGGGLKLKAGKKNLFKVGKEKKKKNKDDKEKIDPDTVENGGWRKIADEFDMKGGTNVAIEV
Consensus      A:YS VK.  L LK. K :     .:.KK:::::.:: . D .  G W .: :  ::.G  .:AIE:

                                                Fascin-like domain
H. sapiens    DKGT-----YIHALDNGLFTLGAPHKEVDEGPSPPEQFTAVKLSD-SRIALKSGYGKYLGINSDGLVVG
M. musculus   DKGA-----YIHALDNGLFTLGAPHREVDEGPSPPEQFTAVKLSD-SRIALKSGYGKYLGINSDGLVVG
X. laevis     DKGA-----YIHALDNGLFTIGAPHKDDDDGPSPPEQFTAIKLSD-SRVALKSGYGKYLGINSDGLVIG
C. elegans    ASGAGSTRTYVAAMDNGKFTIGFPHPE-GEGPNPEEIFALVKTPDDSKISLKTGFGRYVGVDSEYQLVA
Consensus      .G:     Y: A:DNG FT:G PH : .:GP.P E F: :K .D S:::LK:G:G:Y:G::S:  ::.

                                Fascin-like domain
H. sapiens    RSDAIGPREQWEPVFQNGKMALLASNS-CFIR--CNEAGDIEAKSKTAGEEEMIKIRSCAERETKKKDD
M. musculus   RSDAIGPREQWEPVFQDGKMALLASNS-CFIR--CNEAGDIEAKNKTAGEEEMIKIRSCAERETKKKDD
X. laevis     RSDAIGAREQWEPVFDTGKMALLASNS-CFVG--CNEEGDLVAQSKTAGEGEMIKIRSCAEREAKRDDD
C. elegans    MAEAIGSREQFVLVFQEGKTAFQAVSSPLFLSTVPNKEGHIYVASRTATENEMVNIRTDAIQEGP-VDW
Consensus      ::AIG.REQ:  VF: GK A: A .S  F:    N: G.: . .:TA E EM::IR: A :E    D

                                                                Bipartite NLS
H. sapiens    IPEEDKGNVKQCEINYVKKFQS----FQDHKLKISKEDSKILKKARKDGFLHETLLDRRAKLKADRYCK
M. musculus   IPEEDKGSVKQCEINYVKKFQS----FQDHKLKISKEDSKILKKARKDGFLHETLLDRRAKLKADRYCK
X. laevis     IPNEDKGNVKQCEINYVKKFQS----FQDRKLKVSREDNKDLKKARKDGNFHETLLDRRAKLKADRYCK
C. elegans    RSVEDRKNARECETAYVKMYQHSKVDLKNRHIAIDVKDKKGVKKAQADGSAHELLLDRRMKMKSDRYC-
Consensus      . ED: ..::CE  YVK :Q     :::::: :. :D.K :KKA: DG  HE LLDRR K:K:DRYC
```

Figure 1.5: FRG1 protein is very highly conserved in vertebrate and invertebrate. A ClustalW sequence comparison of the predicted amino acid sequences for the human, mouse, *Xenopus laevis*, *Drosophila melanogaster*, and *Caenorhabditis elegans* FRG1 homologs.

CHAPTER 2*

SUBCELLULAR LOCALIZATIONS OF ENDOGENOUS FRG1

* Some of the data and writing is adapted from

Meredith L. Hanel, Chia-Yun Jessica Sun, Steven W. Long, Takako I. Jones, Derek Milner, and Peter L. Jones. ***Facioscapulohumeral muscular dystrophy (FSHD) region gene 1 (FRG1) is a dynamic nuclear and sarcomeric protein.*** Submitted for publication.

INTRODUCTION

FRG1, located 125kb centromeric to the FSHD1A deletion, was one of the early candidate genes for FSHD (van Deutekom et al., 1996). However, recent expression studies have failed to find significant FRG1 misexpression in numerous FSHD patient derived muscle cells and biopsies casting doubt on its involvement in mediating FSHD pathology (Gabellini et al., 2002; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; van Deutekom et al., 1996). Complicating the issue is the lack of understanding towards FRG1's cellular function and it's potential role during normal human muscle development. Initial studies using *Xenopus* as a model for vertebrate development found frg1 was widely expressed early and throughout development, showing elevated levels in vascular tissues and developing muscles with preferential expression in the capillaries, veins, and arteries located between muscle fibers (Hanel et al., 2009; Wuebbles et al., 2009). Knockdown and overexpression experiments confirmed a necessary role for frg1 in development of the musculature and vasculature. Systemic increases in frg1 levels had specific effects on the developing musculature and vasculature, impairing myogenesis and muscle precursor cell migration and causing spurious angiogenesis leading to a tortuous vasculature (Hanel et al., 2009; Wuebbles et al., 2009). These phenotypes are consistent with the two major pathologies seen in FSHD

patients (Gieron et al., 1985; Padberg, 1982). A similar developmental analysis of the *C. elegans* FRG1 ortholog (FRG-1) showed the conserved expression profile of very early, throughout development, and in many different cell-types (Liu et al., 2010). Again, the development, organization, and integrity of the adult body wall musculature were unique in their susceptibility to increased FRG-1 levels. Interestingly, FRG-1 had to be overexpressed in the spaciotemporal pattern dictated by the FRG-1 promoter and there was no affect on the musculature when FRG-1 was overexpressed specifically in adult muscle. Thus, although FRG1 may function in many tissues, the developing musculature and vasculature are uniquely susceptible to systemic changes in FRG1 levels suggesting FRG1 has tissue specific functions. In FSHD, pathogenic changes in FRG1 expression may be occurring early in muscle development or also involve non-myogenic cell lineages (Hanel et al., 2009; Liu et al., 2010; Wuebbles et al., 2009).

Little is known about FRG1's precise function. Overexpression studies in cell culture initially characterized FRG1 as a nuclear and predominantly nucleolar protein (van Koningsbruggen et al., 2004; van Koningsbruggen et al., 2007). In addition, FRG1 has been identified as a component of the spliceosome (Rappsilber et al., 2002). Therefore, FRG1's biological activity likely involves aspects of RNA biogenesis. Our previous work in *C. elegans* showed that the endogenous FRG-1 is both a nuclear and cytoplasmic protein, localizing to the nucleoli and body wall muscle dense bodies, respectively (Liu et al., 2010). Consistent with its localization to muscle attachment sites, FRG-1 was shown to exhibit F-actin binding and bundling activity and this activity was conserved with its human homolog, FRG1 (Liu et al., 2010). This is particularly intriguing from a muscular dystrophy perspective since the *C. elegans* dense bodies form the muscle attachments and

function analogous to the vertebrate Z-discs and costameres combined [reviewed in(Moerman and Williams, 2006)]. Many proteins within or associated with the structures involved in the contractile apparatus and the muscle attachment structures are linked to other myopathies [reviewed in (McNally and Pytel, 2007; Selcen and Carpen, 2008)]. While providing potential insight into FRG-1's function in human muscle development and disease, it is not known how these results translate to the human condition and potentially FSHD. Here, we present an analysis of endogenous human FRG1 in muscle cells during myogenesis and in mouse myofibers. We identified three subcellular pools of FRG1, nuclear, cytoplasmic, and sarcomeric in muscle and show that FRG1 is dynamic, shuttling between the nucleus and cytoplasm. Interestingly we found FRG1's subcellular distribution during human skeletal muscle myogenesis changes dramatically, finally becoming sarcomeric. Overall, our results show that in vertebrates FRG1 is a dynamic nuclear and sarcomere-associated protein suggesting linked roles in the nucleus and at muscle attachment sites.

RESULTS

Endogenous FRG1 has nuclear and cytoplasmic subcellular localizations

To understand the potential for FRG1 to mediate FSHD pathophysiology we first need to understand the normal cellular and developmental function of FRG1 in human muscle. We have recently characterized the *C. elegans* FRG1 homolog as being both a nuclear protein and associated with body-wall muscle sarcomeres (Liu et al., 2010). To characterize the endogenous human FRG1 protein in respect to subcellular localization in skeletal myoblasts and through myogenesis into myotubes, we generated three

independent highly specific anti-FRG1 antibodies, HS1, HS2, and DM1 (Figure 2.1). Western blotting of HeLa whole cell extracts showed each affinity-purified antibody reacted to the predicted 30kDa FRG1 polypeptide, however, the DM1 antibody also recognized a smaller 18kDa polypeptide (Figure 2.2M). The identity of this smaller fragment is not known, however, BLAST searches of the NCBI human non-redundant protein database indicated that the peptides used for immunization are each unique in the human genome and the 18kDa polypeptide likely represents an alternative form of FRG1. The 4q35 localized FRG1 gene, the only locus containing all nine exons, is partially duplicated at several other genomic loci and alternatively spliced FRG1 transcripts have been previously reported (van Deutekom et al., 1996). Although all three antibodies appear highly specific for FRG1 by western blotting, the primary technique used in this study is immunocytochemistry (ICC). Therefore, to characterize the specificity of the antibodies for ICC, a specific siRNA-mediated knockdown approach was used. HeLa cells were transfected three consecutive times with a pool of 4 siRNAs specific to human FRG1, subjected to immunostaining with each of the FRG1 antibodies, and compared to controls (Figure 2.2A-L). Each of the FRG1 antibodies showed similar immunostaining patterns with varying intensities of both cytoplasmic and nuclear staining. In all cases the FRG1 signal was severely depleted or absent in the siRNA treated cells (Figure2.2, compare C, D with A, B; G, H with E, F; and K, L with I, J). Due to the fact that transfection efficiency was less than 100%, some cells were still positive for FRG1 staining and served as positive controls for the immunostaining procedure. We conclude that all three of our FRG1 antibodies are specific for FRG1 in immunohistochemistry.

As opposed to overexpressed epitope-tagged FRG1 that appears exclusively nuclear in cell culture (van Koningsbruggen et al., 2004), the endogenous FRG1 in HeLa cells clearly appears to be both nuclear and cytoplasmic by immunostaining. Western blotting of HeLa cell extract fractionated into nuclear and cytoplasmic pools confirmed the dual subcellular localization of FRG1 (Figure 2.2N). Although HeLa cells are a commonly used human cell line for many studies of protein function because they are fast growing and easily transfected, they are HPV17 transformed immortal adenocarcinoma cells, quite distant from the muscle cells in which we are interested in studying. Therefore, the unexpected cytoplasmic subcellular localization of human FRG1 was further addressed in multiple cell types including HeLa cells, normal human skeletal myoblasts, murine muscle derived stem cells (MDSC), and murine C2C12 cells (Figure 2.3). In HeLa and C2C12 cells, FRG1 was predominantly localized to nuclei, however the cells displayed distinct, non-uniform fiber-like cytoplasmic FRG1 immunostaining strongly suggesting an association with a subcellular architecture. In the human myoblasts (Figure 2.3Q) and the murine MDSCs (Figure 2.3U) the nuclear FRG1 was much less pronounced compared with the cytoplasmic staining which appeared to surround the nucleus and was very granular. We conclude that endogenous FRG1 exists in both a nuclear and cytoplasmic pool in all cell types tested, however its cytoplasmic to nuclear distribution is cell type dependent.

FRG1 is a nuclear-cytoplasmic shuttling protein

The endogenous FRG1 is localized in both the nucleus and cytoplasm. Nuclear shuttling assays were performed (Figure 2.4) to determine if these two pools of FRG1

were linked. Murine C2C12 cells, easily identifiable by their DNA-dense nuclear foci, were transfected with a plasmid expressing epitope tagged HA-FRG1 and allowed to accumulate HA-FRG1 overnight. Cycloheximide (CHX) was added to the culture media to block translation and the cells were fused with non-transfected HeLa cells, readily identifiable by their DNA poor nucleoli, in continued presence of CHX, and HA-FRG1 localization was monitored over time by ICC probing for HA. Thus, any HA signal in the HeLa cells represents FRG1 protein synthesized in the C2C12 cells. Within two hours of starting the fusion FRG1 synthesized in a C2C12 cell (Figure 2.4A-D, white arrow) had begun to accumulate in the nuclei and concentrate the in nucleoli of a fused HeLa cell (Figure 2.4A-D, blue arrow). This nuclear import of FRG1 was more evident at three hours (Figure 2.4E-H) and at four hours appeared to have reached equilibrium (Figure 2.4I-L). As the amount of cytoplasmic HA-FRG1 is almost undetectable, we deduce that much of the HeLa nuclear HA-FRG1 came from the C2C12 nuclear FRG1 and conclude that FRG1 shuttles between the nucleus and cytoplasm.

FRG1's subcellular localization changes during myogenic differentiation

FRG1 is critical for development of the musculature and the vasculature (Hanel et al., 2009; Wuebbles et al., 2009); therefore, the newly described subcellular dynamics for the endogenous FRG1 were examined during myogenesis (Figure 2.5). Normal skeletal muscle myoblasts were stimulated to undergo differentiation by serum depletion and analyzed by ICC at various time points to determine FRG1's subcellular distribution. In undifferentiated primary myoblasts, FRG1 was almost exclusively cytoplasmic (Figure 2.5A, B), however within 24 hours of beginning differentiation FRG1 became

predominantly nuclear, and strongly nucleolar (Figure 2.5C, D), and by five days post-differentiation the majority of FRG1 appeared to be predominantly cytoplasmic again, although now appearing in stress-fiber-like structures (Figure 2.5E, F). Interestingly, by eight days post-differentiation the cytoplasmic FRG1 appeared in a striated pattern reminiscent of sarcomeres with the immunostaining being consistent among all three FRG1 antibodies (Figure 2.5G, J, M). Co-staining for sarcomeric α-actinin showed clear co-localization of the striated FRG1 signal with α-actinin (Figure 2.5G-O, arrows), indicating that FRG1 was in fact localizing to the sarcomeric structures. Considering that from 2-5 days post-differentiation, α-actinin was aggregating at the sarcolemma (Figure 2.5D, F), forming Z-disc-like structures in the absence of any detectable localized FRG1, we conclude that FRG1 associates with more matured Z-discs, after adjacent myofibrils are aligned, and is not likely involved in their establishment or initial assembly.

Sarcomeric FRG1 is associated with the Z-disc

To determine if FRG1 was a new sarcomere-associated protein, intact mouse myofibers were immunostained for FRG1, sarcomeric α-actinin, and desmin (Figure 2.6). All three FRG1-specific antibodies showed intense striated patterns of FRG1 (Figure 2.6A, E, I, M, Q) as well as some nuclear staining (Figure 2.6I, Q, blue arrows) with the striated FRG1 immunostaining overlapping with sarcomeric α-actinin (Figure 2.6D, H, L). Higher resolution confocal images revealed the precise co-localization of FRG1 with α-actinin (Figure 2.6M-P) however FRG1 and desmin, while displaying highly similar patterns, did not precisely co-localize (Figure 2.6Q-T). Negative controls using secondary antibodies alone showed no signal (data not shown). To further characterize the

sarcomeric FRG1, purified myofibrils were immunostained for FRG1, using rhodamine-phalloidin as a marker for the Z-disc. Here, FRG1 showed a much more diffuse pattern (Figure 2.6U-W) than seen on the intact myofibers suggesting FRG1 is less stably associated with the individual myofibrils than the Z-discs of intact myofibers.

DISCUSSION

Expression analyses have failed to produce consistent, reproducible results showing any 4q35 FSHD candidate gene, including FRG1, is misexpressed in FSHD muscle biopsies or patient-derived myocytes (Gabellini et al., 2002; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007). An alternative approach using overexpression of FSHD candidate genes in animal models has singled out FRG1 alone as being able to recapitulate both muscular and vascular FSHD-like pathology when overexpressed systemically (Gabellini et al., 2006; Hanel et al., 2009; Liu et al., 2010; Wuebbles et al., 2009). However, these models have been criticized for exaggerated levels of expression beyond what would be expected in FSHD and therefore potentially leading to artificial phenotypes resulting in an inconclusive verdict. To gain further insight on the viability of FRG1 misexpression being involved in FSHD, we have sought here to further understand the endogenous FRG1 protein in mouse muscle and differentiating myotubes derived from human skeletal myoblasts. Although FRG1 is ubiquitously expressed in all tissues tested by mRNA analysis (van Deutekom et al., 1996), our ICC analyses showed the FRG1 protein is specifically localized within myotubes and myofibers (Figure 2.5 and 2.6). Both aspects are interesting from muscle development and muscular dystrophy perspectives.

Previous cell culture studies using epitope-tagged FRG1 transgenes characterized FRG1 as near exclusively nuclear with strong nucleolar and nuclear speckle concentrations implicating FRG1 in RNA biogenesis (van Koningsbruggen et al., 2004; van Koningsbruggen et al., 2007). Although our analysis of the endogenous FRG1 in cell culture and myofibers clearly contradict the characterization of an exclusive nuclear localization for FRG1 (Figure 2.2, 2.3, 2.5, and 2.6), endogenous FRG1 does accumulate in the nucleoli during myotube formation supporting the claim for a role in RNA biogenesis. It should be noted than in our nuclear shuttling assays, HA-FRG1 recipient cells (HeLa) accumulated the transiently expressed FRG1 almost exclusively in their nuclei and specifically in the nucleoli (Figure 2.4), despite the endogenous FRG1 showing both cytoplasmic and nuclear staining (Figure 2.2 and 2.3). This data indicates that the majority of overexpressed FRG1 protein is nuclear and preferentially nucleolar. Thus, this raises the question of how is the exogenous or overexpressed FRG1 different from the endogenously regulated FRG1? It is interesting to note that different cell types showed different ratios of nuclear to cytoplasmic FRG1 with undifferentiated and fully differentiated muscle cells showing the greatest amount in the cytoplasm. Since exogenous or overexpressed FRG1 preferentially accumulates in the nucleus, potentially a certain cell-type specific level of endogenous FRG1 is capable of being actively maintained in the cytoplasm (FRG1 is ~29kDa) at any one time and any increases in FRG1 protein levels result in default FRG1 nuclear localization. Combined with the dynamics shown by the nuclear shuttling assays, this would predict that eventually the exogenous or overexpressed FRG1 would show the similar cytoplasmic staining albeit with the intense nuclear staining, dependent on the turnover rate for the endogenous

cytoplasmic retained FRG1. This is in fact exactly what was seen in the *C. elegans* study on FRG-1; the overexpressed epitope tagged FRG-1 intensely localized to the nuclei yet over time, faint but detectable FRG-1 localization was seen in the body-wall muscle attachment sites (Liu et al., 2010). We suggest that in our nuclear shuttling assay and published overexpression studies, the overexpressed FRG1 is actively shuttling between the nucleus and cytoplasm but is visualized exclusively in the nuclei because it is not being readily retained in the cytoplasm. Conversely, the endogenous FRG1 is stably maintained in the cytoplasm awaiting a signal to release it to the nucleus. This cytoplasmic retention model is supported by the dramatic change in endogenous FRG1 localization to the nucleus in myoblasts upon stimulation of myogenic differentiation (Figure 2.5). This model further predicts that even small changes in FRG1 levels would alter its subcellular distribution, aberrantly increasing its levels in the nucleus.

Active cytoplasmic retention of FRG1 likely involves interaction with other proteins to anchor it. Recently we showed that FRG1 is a bona fide F-actin binding and bundling protein (Liu et al., 2010), further supporting a cytoplasmic role for FRG1. Here we have characterized FRG1 as a sarcomeric Z-disc associated protein in mouse and myotubes derived from human skeletal myoblasts (Figure 2.5 and 2.6). This highlights that FRG1, although ubiquitously expressed in respect to tissues, has cell type specific, and particularly muscle specific, functions. If FRG1 were misregulated in FSHD, this could explain why the genetic lesion presents skeletal muscle specific pathophysiology. The Z-disc localization is additionally intriguing in respect to FSHD since in FSHD patient muscle some of the structures at the sarcolemma are misaligned and the association of the sarcolemma with contractile structures is altered (Reed et al., 2006).

Numerous other myopathies can trace their molecular defects to proteins associated with the contractile apparatus and force generating structures of skeletal muscle (McNally and Pytel, 2007; Selcen, 2008; Selcen and Carpen, 2008). This work places FRG1 as the only current FSHD candidate gene whose product is directly linked to the skeletal muscle contractile apparatus.

Identifying FRG1 as a dynamic nuclear and sarcomere-associated protein may suggest a link between the two known biological activities for FRG1, F-actin binding/bundling and RNA biogenesis (Liu et al., 2010; van Koningsbruggen et al., 2004; van Koningsbruggen et al., 2007). Potentially FRG1 could be transducing signals from the Z-disc to the nucleus directly and affecting mRNA biogenesis, as is the case for the dynamic Z-disc protein MLP (Arber et al., 1994; Knoll et al., 2002). Conversely, FRG1 may be involved in trafficking molecules such as RNAs to the Z-disc for site-specific translation. In either case, one can imagine how aberrantly altering the levels of FRG1, and thus affecting FRG1-mediated signaling or transport, could disrupt the efficiency of myogenic differentiation, muscle regeneration, and maintenance of muscle integrity over time as seen in the animal models overexpressing or depleting FRG1, and as proposed for FSHD (Wuebbles et al., 2009).

Overall, we identify FRG1 as a dynamic nuclear and cytoplasmic protein associated with mature sarcomeres in mouse skeletal muscle and myotubes derived from human skeletal myoblasts. Thus, FRG1, an actin-bundling protein previously shown to be critical for muscle and vascular development, provides the only link between a FSHD candidate gene and the muscle contractile machinery.

MATERIALS AND METHODS

Cell culture

HeLa cells and C2C12 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS) 2 mM L-glutamine, and 1% penicillin–streptomycin (pen/strep). Proliferating primary human skeletal muscle myoblasts (HSMM) were obtained from Lonza (Walkersville, MD) and were seeded on 0.02% collagen-coated surfaces and maintained in SkBM-2 medium supplemented with SkGM-2 SingleQuots (Lonza) according to the manufacturer's instructions. For myotube formation, HSMMs were seeded on 0.02% collagen-coated coverslips at 1.5 x $10^4/cm^2$ density for ICC analysis, and the following day were induced to form myotubes by adding fusion medium (DMEM/F-12 50:50 supplemented with 2% horse serum). Murine muscle derived stem cells (MDSC) were isolated and cultured as described (Qu-Petersen et al., 2002). For ICC analysis, MDSCs were seeded on 0.02% collagen-coated coverslips.

Preparation of myofibers

Mice (C57/B6) were humanely euthanized in accordance with approved UIUC IACUC protocols. Adult mouse muscle fibers were isolated from the flexor digitorum brevis muscle of 1-3 month old female mice. Isolated muscles were washed briefly in DMEM, then incubated in DMEM with 0.2% collagenase type I (Worthington Biochemical, Lakewood, NJ) for 4 hrs at 37°C, with gentle agitation every 15 min and changes into fresh collagenase solution every hour. At the completion of digestion, excess tendon material was carefully dissected away, and fiber bundles were transferred to a dish of myofiber medium (DMEM supplemented with 10mM HEPES, 5% heat-inactivated

horse serum, 1% pen/strep and 0.1% amphotericin B). Individual fibers were liberated from the muscle mass by gentle tituration and agitation, and cultured overnight in myofiber medium at 37°C, 5%CO_2. The following day, healthy, undamaged myofibers were plated on glass coverslips coated with Geltrex (Invitrogen, Carlsbad, CA) and allowed to adhere for 1-2 hrs before fixation.

Preparation of myofibrils

Myofibrils were prepared as described (Qu-Petersen et al., 2002). Skeletal muscle was dissected from rectus femoris muscle, teased into thin 3-5 mm strips, and soaked in EGTA-Ringer's solution (100 mM NaCl, 6 mM KH2PO_4, 2 mM KCL, 2 mM $MgCl_2$, 1 mM EGTA, 0.1% glucose, pH 7.0 at 0°C) overnight at 4°C. Muscle strips were transferred to ice-cold rigor buffer (100 mM KCl, 2 mM $MgCl_2$, 10 mM KH_2PO_4, 1 mM EDTA, 0.5 mM DTT, pH 7.0 at 0°C) and homogenized using a loose pestle Dounce tissue grinder until a majority of individual myofibrils were obtained by observing under a phase contrast microscope. The homogenate was centrifuged at 200xg for 5 min to remove debris and the supernatant was centrifuged at 2000xg for 5 min to pellet myofibrils. The myofibrils were washed 5 times with ice-cold rigor buffer, plated on coverslips coated with poly-L-lysine, and immediately fixed for immunofluorescence staining.

Protein extracts

To generate whole cell extracts (WCE), cells were collected in PBS, pelleted, resuspended in 10 cell volumes of RIPA+ buffer (150 mM NaCl, 1% IGEPAL, 0.5% sodium deoxycholate, 50 mM Tris pH 8.0, + 1% SDS), and incubated on ice for 10 min.

Lysed cells were sonicated briefly, centrifuged at 100,000xg for 10 min and the soluble fraction was used for western blotting analysis. Nuclear and cytoplasmic fractions were generated as described (Shapiro et al., 1988). The PCNA rabbit monoclonal antibody (Epitomics, Inc, 2755-1) was used as marker for the nuclear fraction.

FRG1 antibodies

The HS1, HS2, and DM1 rabbit polyclonal antibodies were made by GenScript USA Inc. (Piscataway, NJ) and generated against synthesized peptide antigens conjugated to KLH. The HS1 peptide (KKDDIPEEDKGNVK) and HS2 peptide (GRSDAIGPREQWEP) were from the predicted human FRG1 amino acid sequence while the DM1 peptide (TLLDRRSKMKADRYC) was from the predicted *Drosophila melanogaster* FRG1 (CG6480) amino acid sequence (Figure 2.1). Antisera were affinity purified against the peptide cross-linked to NHS-Sepharose (GE Healthcare), eluted in 10 mM glycine, pH 2.5, and dialyzed against PBS pH7.4.

FRG1 siRNA knockdown

The On-Target plus SMARTpool siRNA reagent for human FRG1 (Dharmacon, Inc) containing four duplex siRNAs (siRNA1: GUUUACGGCUGUCAAAUUA; siRNA2: CGACAGAUACUGCAAGUGA; siRNA3: GGAACCAAGACGAAGAGUA; siRNA4: AAACCCAGCUUGAUAUUGU) was used for FRG1 knockdown in cell culture. Transfections were carried out using oligofectamine (Invitrogen) at a final concentration of 100 nM siRNAs. HeLa cells (~50% confluent) were transfected three times in 24 hr intervals and then assayed for FRG1 by ICC.

Nuclear shuttling assay

The assay was carried out essentially as described (Kawamura et al., 2002). The HA-FRG1 expression plasmid was generated by subcloning the human FRG1 coding sequence into pcDNA3.1 HA (Matzat et al., 2008). Murine C2C12 cells (~60% confluent) were transfected with pcDNA3.1HA-FRG1 using TransIT-LT1 reagent (Mirus Bio LLC, Madison, WI) and allowed to grow for 24 hrs. The cells were removed by trypsinization, washed with PBS, plated on poly-L-lysine coated coverslips ($1 \times 10^5/cm^2$) and allowed to adhere for 2 hrs before non-transfected HeLa cells were overlaid ($5 \times 10^4/cm^2$) onto the transfected C2C12 cells for 3 hours. The co-cultures were incubated with 100mg/ml Cycloheximide (CHX) for 15 min to stop translation, and the cells were fused by adding 50% (wt/vol) polyethylene glycol 4000 in DMEM for 2 min. The fusions were immediately washed with DMEM and then incubated with 100mg/ml CHX for 2, 3, or 4 hrs followed by ICC analysis. The cells were fixed with 4% formaldehyde (FA) in PBS for 15 min, immunostained with HA monoclonal antibody clone 3F10 (1:100)(Roche) as described below, and co-stained stained with Hoechst 33342 (5 mg/ml in PBS).

ICC staining

HeLa, C2C12, and MDSC, were fixed in 4% FA in PBS and HSMM were fixed in 2% FA in PBS, for 15 min at room temperature (RT). After fixation, cells were permeabilized with 0.25% Triton-X 100 in PBS for 10 min on ice, and subsequently blocked with 2% BSA in PBS for 30 min at RT. Primary antibody incubations were carried out at RT for 1 hr up to overnight at 4°C and secondary antibody incubations were for 40 min at RT. Mouse myofibers were fixed in 2% paraformaldehyde for 15 minutes,

rinsed with PBS and permeabilized with 0.1% Triton X-100 for 10 min. For blocking, fibers were incubated in TBS-T +5% milk powder and 0.02% sodium azide for 1 hr at RT, or alternatively, overnight at 4^{o}C, and then incubated with diluted primary antibodies and secondary antibodies as above. Mouse myofibrils were fixed with 2% FA in PBS for 15 min at RT, rinsed with PBS and permeabilized with 0.1% Triton X-100 for 10 min. Myofibrils were blocked with normal goat serum for 30 min at RT then incubated with FRG1 primary antibody followed by Alexa 488-conjugated goat anti-rabbit IgG secondary antibody as described above. The antibodies and their dilutions were as follows: desmin monoclonal antibody [D9] (Santa Cruz Biotechnology Inc, sc-52326), 1:1000; sarcomeric α-actinin mouse monoclonal antibody [EA-53] (Abcam Inc, ab-9465), 1:200; FRG1 HS1 (1:100), HS2 (1:200), and DM1(1:200). Secondary antibodies used were FITC-conjugated goat anti-mouse, FITC-conjugated donkey anti-rabbit (pre-cleared), and rhodamine-conjugated goat anti-rabbit (Jackson ImmunoResearch Laboratories Inc) used at 1:100. Alexa488-conjugated goat anti-rabbit IgG, highly absorbed (Invitrogen) used at 1:800. F-actin was visualized with 5 units/ml rhodamine-phalloidin (Invitrogen) incubated for 30 min at RT for cell culture staining, and with 1 unit/ml rhodamine-phalloidin incubated for 2 min at RT for myofibrils. DAPI was used at 1μg/ml.

Microscopy

Fluorescence mages were taken by fluorescence microscopy using an Olympus BX60 microscope equipped with a SpotRT monochrome model 2.1.1 camera and Spot Advanced software (Diagnostic Instruments, Sterling Heights, MI). Confocal microscopy was carried out with Zeiss LSM510. Immunohistochemical images were acquired with

Olympus BX60 microscope equipped with a Leica DFC290 camera using Leica Application Suit software (Leica microsystems). All images were processed using Adobe Photoshop to adjust brightness, contrast, size, and merged or split channels.

REFERENCES

Arber, S., G. Halder, et al. 1994. Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation. Cell, **79**: 221-31.

Gabellini, D., G. D'Antona, et al. 2006. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. Nature, **439**: 973-7.

Gabellini, D., M. R. Green, et al. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell, **110**: 339-48.

Gieron, M. A., J. K. Korthals, et al. 1985. Facioscapulohumeral dystrophy with cochlear hearing loss and tortuosity of retinal vessels. Am J Med Genet, **22**: 143-7.

Hanel, M. L., R. D. Wuebbles, et al. 2009. Muscular dystrophy candidate gene FRG1 is critical for muscle development. Dev Dyn, **238**: 1502-12.

Jiang, G., F. Yang, et al. 2003. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet, **12**: 2909-21.

Kawamura, H., Y. Tomozoe, et al. 2002. Identification of the nucleocytoplasmic shuttling sequence of heterogeneous nuclear ribonucleoprotein D-like protein JKTBP and its interaction with mRNA. J Biol Chem, **277**: 2732-9.

Klooster, R., K. Straasheijm, et al. 2009. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. Eur J Hum Genet, **17**: 1615-24.

Knoll, R., M. Hoshijima, et al. 2002. The cardiac mechanical stretch sensor machinery involves a Z disc complex that is defective in a subset of human dilated cardiomyopathy. Cell, **111**: 943-55.

Liu, Q., T. I. Jones, et al. 2010. Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. J Cell Sci, **123**: 1116-23.

Masny, P. S., O. Y. Chan, et al. 2010. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. Eur J Hum Genet, **18**: 448-56.

Matzat, L. H., S. Berberoglu, et al. 2008. Formation of a Tap/NXF1 homotypic complex is mediated through the amino-terminal domain of Tap and enhances interaction with nucleoporins. Mol Biol Cell, **19**: 327-38.

McNally, E. M. and P. Pytel. 2007. Muscle diseases: the muscular dystrophies. Annu Rev Pathol, **2**: 87-109.

Moerman, D. G. and B. D. Williams. 2006. Sarcomere assembly in C. elegans muscle. WormBook: 1-16.

Osborne, R. J., S. Welle, et al. 2007. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology, **68**: 569-77.

Padberg, G. W. 1982. Facioscapulohumeral Disease.

Qu-Petersen, Z., B. Deasy, et al. 2002. Identification of a novel population of muscle stem cells in mice: potential for muscle regeneration. J Cell Biol, **157**: 851-64.

Rappsilber, J., U. Ryder, et al. 2002. Large-scale proteomic analysis of the human spliceosome. Genome Res, **12**: 1231-45.

Reed, P., N. C. Porter, et al. 2006. Sarcolemmal reorganization in facioscapulohumeral muscular dystrophy. Ann Neurol, **59**: 289-97.

Selcen, D. 2008. Myofibrillar myopathies. Curr Opin Neurol, **21**: 585-9.

Selcen, D. and O. Carpen. 2008. The Z-disk diseases. Adv Exp Med Biol, **642**: 116-30.

Shapiro, D. J., P. A. Sharp, et al. 1988. A high-efficiency HeLa cell nuclear transcription extract. DNA, **7**: 47-55.

van Deutekom, J. C., R. J. Lemmers, et al. 1996. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. Hum Mol Genet, **5**: 581-90.

van Koningsbruggen, S., R. W. Dirks, et al. 2004. FRG1P is localised in the nucleolus, Cajal bodies, and speckles. J Med Genet, **41**: e46.

van Koningsbruggen, S., K. R. Straasheijm, et al. 2007. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. Chromosoma, **116**: 53-64.

Wuebbles, R. D., M. L. Hanel, et al. 2009. FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. Dis Model Mech, **2**: 267-74.

FIGURES

Figure 2.1: FRG1 is highly conserved from human to C. elegans. A ClustalW sequence comparison of the predicted amino acid sequences for the human, mouse, *Xenopus laevis*, *Drosophila melanogaster*, and *Caenorhabditis elegans* FRG1 homologs. The location of FRG1 antibodies used in this study are shown and highlighted in yellow. The highlighting indicates species cross-reactivity and amino acid differences are shown in bold red lettering. NLS = nuclear localization signal.

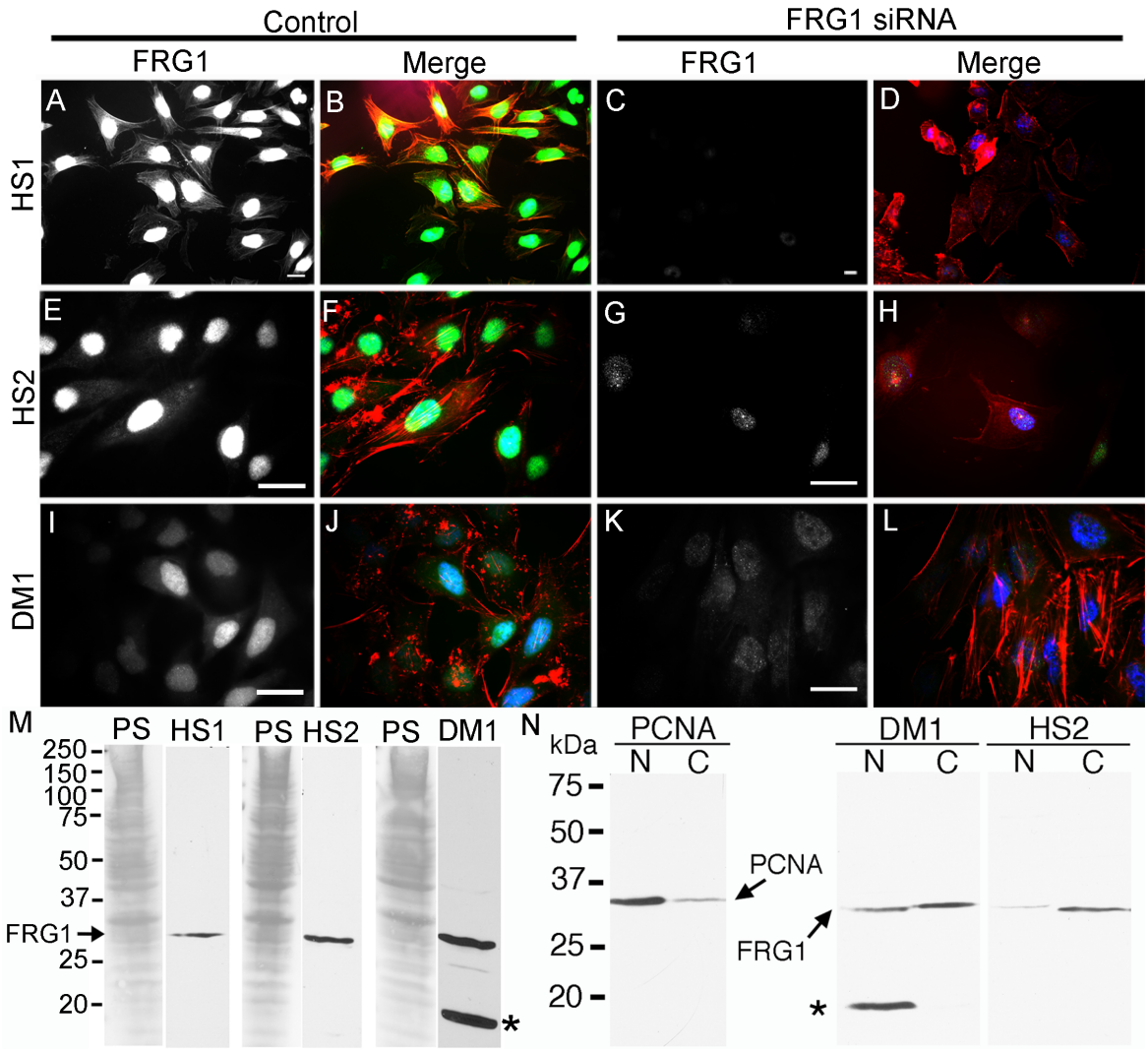

Figure 2.2: FRG1 antibodies are highly specific. Immunocytochemistry on HeLa cells (A, B, E, F, I, J) or HeLa cells transfected with a FRG1-specific pool of siRNAs (C, D, G, H, K, L) using the HS1 (A-D), HS2 (E-H), or DM1 (I-J) FRG1 antibodies show a specific reduction in both the cytoplasmic and nuclear FRG1 antibody signal intensities in the siRNA treated cells. The merged images (B, D, F, H, J, L) show FRG1 (A, C, E, G, I, K) in green, DAPI in blue, and phalloidin in red. Images are taken under the same parameters. Bars = 10 µm. M) Western blot analysis of HeLa whole-cell extract (200 mg/lane) probed with HS1, HS2, and DM1 as indicated. PS indicates Ponceau S staining of membranes. N) Western blot analysis of HeLa cell extract fractionated into nuclear and cytoplasmic pools and probed with DM1 and HS2 as indicated. A monoclonal antibody against proliferating cell nuclear antigen (PCNA) was used to control for contamination of the cytoplasmic protein pool with nuclear proteins.

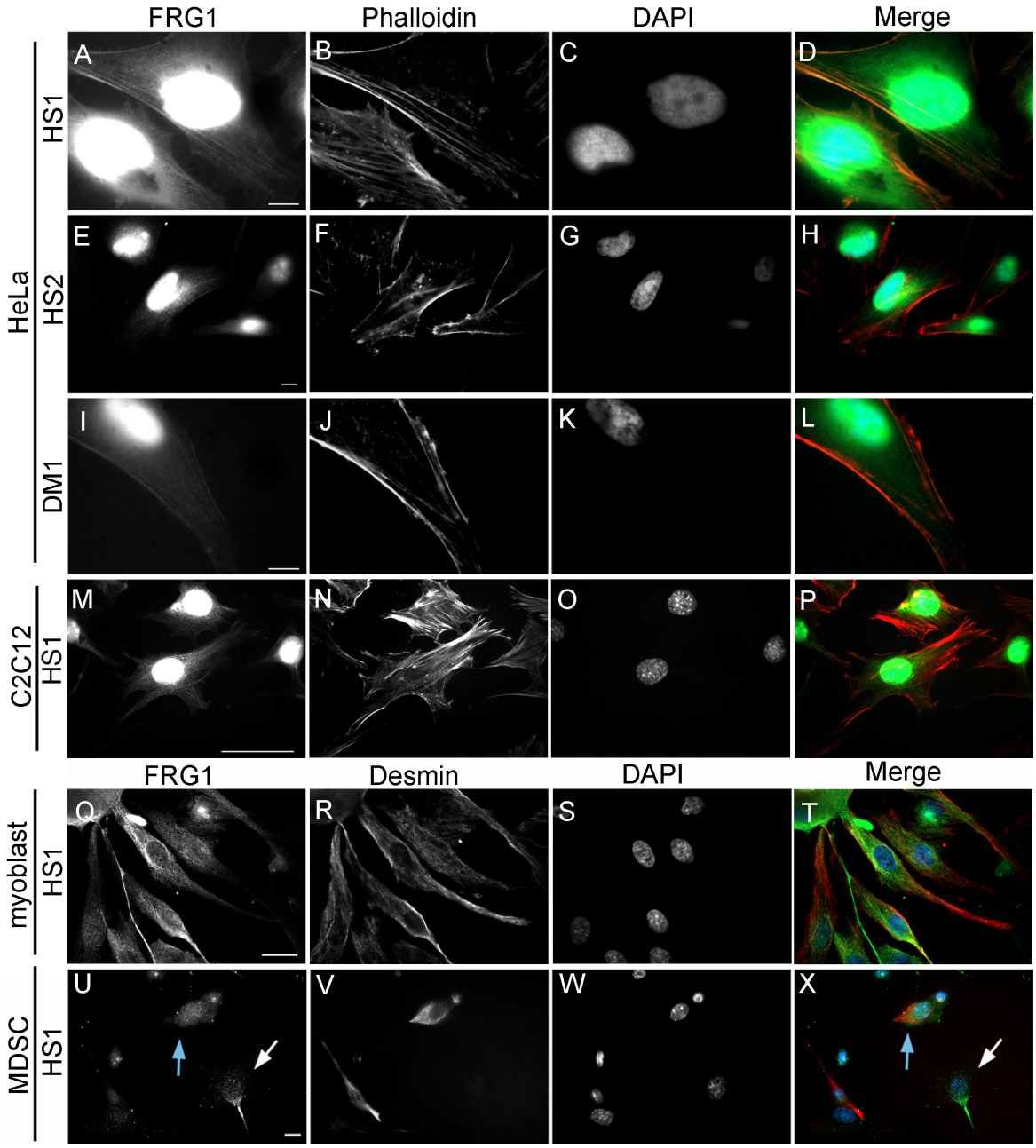

Figure 2.3: Endogenous FRG1 is both a nuclear and cytoplasmic protein in multiple cell types. Immunocytochemistry on cultured (A-L) HeLa cells, and (M-P) murine C2C12 cells reveal intense nuclear FRG1 immunostaining accompanied by cytoplasmic FRG1 immunostaining (A, E, I, M; green in merge). (Q-T) Human skeletal muscle myoblasts, and (U-X) murine muscle-derived stem cells show much more prominent cytoplasmic FRG1 immunostaining (Q, U, T, X; green in merge). Rhodamine-phalloidin staining (B, F, J, N; red in merge) labeled the cytoplasmic actin filaments while DAPI (C, G, K, O, S, W; blue in merge) labeled nuclei. Desmin immunostaining (R, V; red in merge) was used to confirm the myoblast phenotypes.

Figure 2.3 continued: (U-X) White arrow indicates a desmin negative MDSC and blue arrow indicates a desmin positive MDSC beginning differentiation. Overall, the immunostaining patterns within a cell type are consistent between HS1 (A-D, Q-T), HS2 (E-H), and DM1 (I-L, M-P), three independent FRG1 antibodies raised against peptides from different regions of FRG1, as shown in Figure 2.1. Bars = 10 μm.

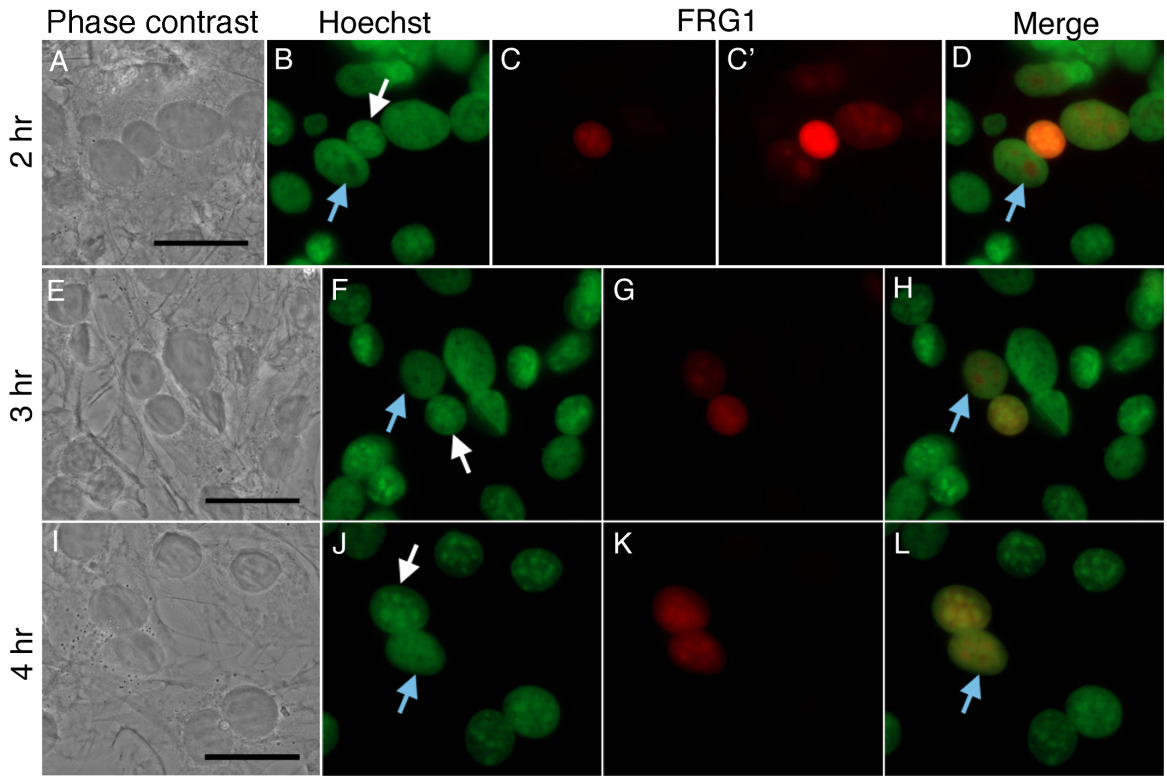

Figure 2.4: FRG1 shuttles between the nucleus and cytoplasm. (A-L) Murine C2C12 cells, morphologically distinguished by their DAPI-bright foci (white arrow), expressing HA-FRG1 (red) and treated with cyclohexamide (CHX) were fused with HeLa cells, distinguished by their DAPI-poor nucleoli (blue arrows) in the presence of CHX. (A-D) Two hours into the fusion process FRG1 translated in the C2C12 cells begins to localize in the HeLa cell nuclei (C', longer exposure of C) and specifically the nucleoli (D, blue arrows). This translocation of FRG1 from C2C12 to HeLa nuclei is more evident at 3 hours (E-H) and at 4 hours (I-L), appearing to have reached equilibrium between the two cell type nuclei (K). Bars = 10 μm.

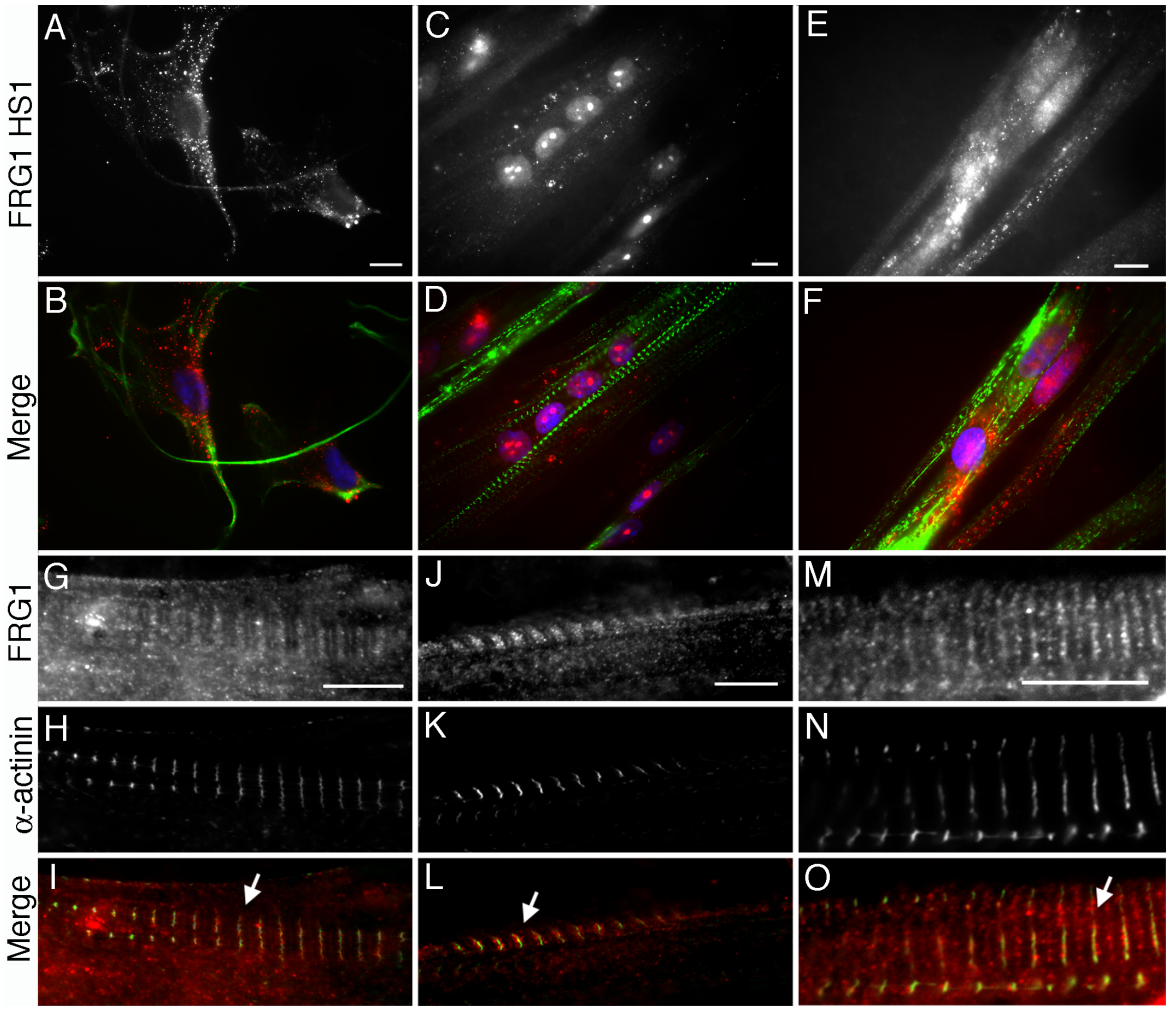

Figure 2.5: FRG1 subcellular localization changes dramatically when primary human skeletal myocytes are stimulated to undergo myogenic differentiation. FRG1 subcellular localization was monitored using the HS1 antibody (A-I red) in normal human skeletal muscle myoblasts (A, B), 2 days post-differentiation (C, D), 5 days post-differentiation (E, F), and 8 days post-differentiation (G-I). Developing Z-discs were monitored by α-actinin immunostaining (D, F, I, L, O; green). FRG1 was similarly monitored with the HS2 (J-L; red) and DM1 (M-O; red) antibodies with only the 8 days post-differentiation shown. Using all three antibodies FRG1 is detected co-localized with α-actinin at Z-discs 8 days post-differentiation, but not earlier (I, L, O; white arrows). Desmin immunostaining (B; green) confirmed the myoblast phenotype and DAPI identified the nuclei (B, D, F; blue). Bars = 10 μm.

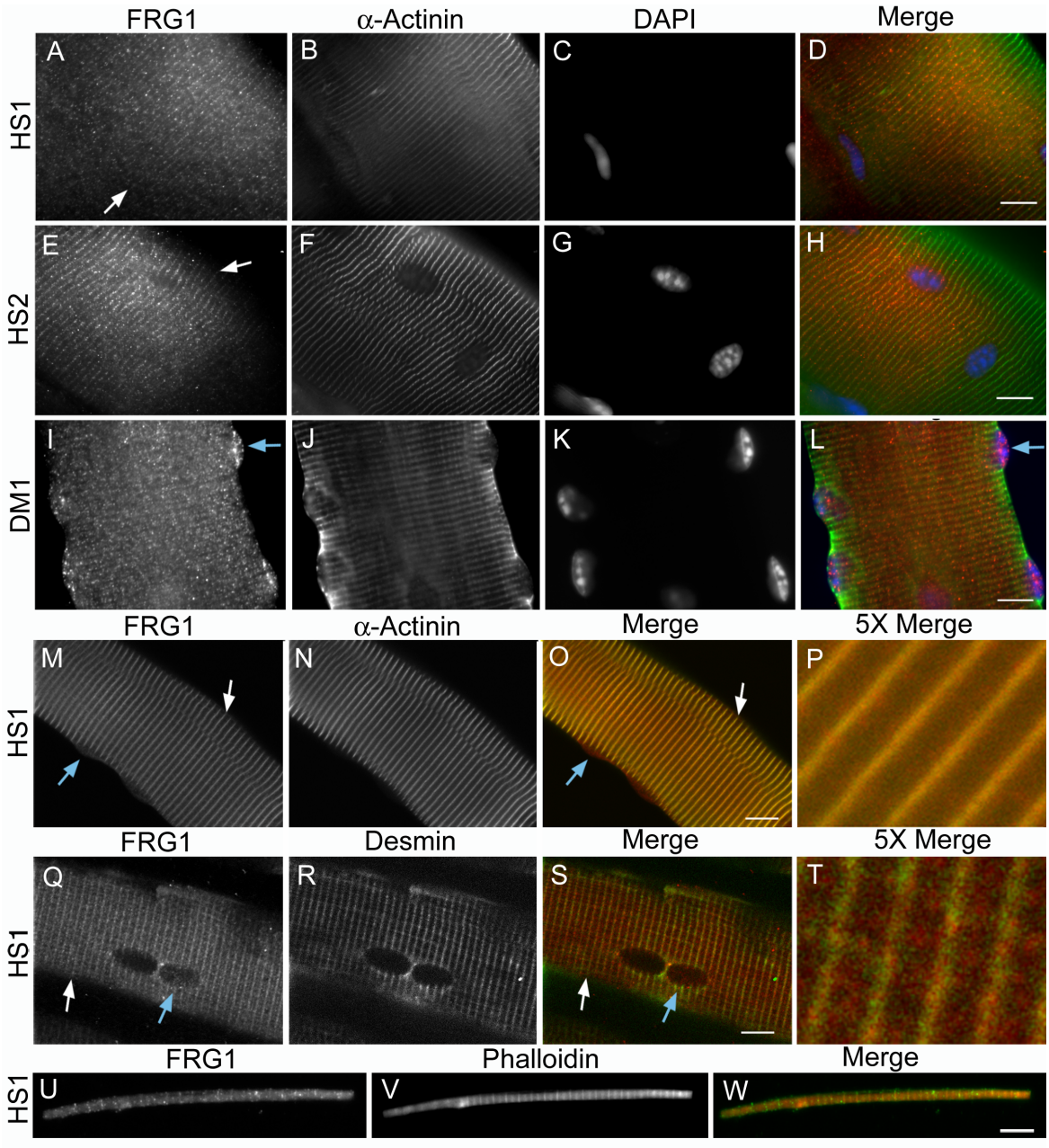

Figure 2.6: FRG1 is a sarcomeric protein localized to myofiber Z-discs. To characterize FRG1's localization at the sarcomere, mouse muscle fibers were isolated and subjected to immunohistochemistry using the HS1 (A, M, Q), HS2 (E) and DM1 (I) FRG1 antibodies and co-staining for α-actinin (B, F, J, N) or desmin (R). Images were visualized by standard fluorescent microscopy (A-L) or by fluorescent confocal microscopy (M-T). When merged (D, H, L,) α-actinin (green) appears overlapping and flanked by the FRG1 (red) around the Z-lines. However, confocal images clearly show the precise co-staining of FRG1 and α-actinin (O, P) while FRG1 and desmin only partially overlap (S, T). In all cases, FRG1 was detectable in myofiber nuclei (blue arrows). Bars = 10 μm. (U-W) Isolated myofibril immunostained with HS1 (U), counter stained with rhodamine-phalloidin (V), and images merged (W). Bar = 10 μm.

CHAPTER 3*

FRG1 IS A RNA-ASSOCIATED PROTEIN

* Some of the data and writing is adapted from

Chia-Yun Jessica Sun, Silvana van Koningsbruggen, Kirsten Straasheijm, Rinse Klooster, Steven W. Long, Takako I. Jones, Michel Bellini, Lyne Levesque, William M. Brieher, Silvère M. van der Maarel, and Peter L. Jones. ***FSHD region gene 1 (FRG1) is a dynamic RNA-associated, actin bundling protein.*** Submitted for publication.

INTRODUCTION

FSHD region gene 1 (FRG1) is a dynamic nuclear and cytoplasmic protein that, in skeletal muscle, shows an additional localization to the sarcomere as described in Chapter 2. FRG1 is required for the normal development of the vertebrate musculature and vasculature and overexpressing FRG1 in vertebrate and invertebrate animal models leads to muscular dystrophy (Gabellini et al., 2006; Hanel et al., 2009; Liu et al., 2010; Wuebbles et al., 2009), however, the precise function of the endogenous human FRG1 is unknown. A putative domain analysis of the predicted FRG1 protein sequence revealed two nuclear localization signals (NLS) residing in the N-terminus, one bipartite NLS in the C-terminus, and a central fascin-like domain, found in the fascin family of actin-bundling and cross-linking proteins (Kureishy et al., 2002). Overall, FRG1 in its entirety is very highly conserved from vertebrates to invertebrate (human FRG1 shares 97%, 81%, and 46% homology with mouse, *Xenopus*, and *C. elegans*, respectively) suggesting important biological functions that were conserved throughout evolution (Grewal et al., 1998).

Several lines of evidence suggest FRG1 is involved in RNA biogenesis. Epitope-tagged FRG1 overexpressed in vertebrate cell culture appears almost completely nuclear, and localized to the nucleoli, Cajal bodies, and nuclear speckles, sites where RNA-

biogenesis is taking place (van Koningsbruggen et al., 2004). Mis-spliced mRNA transcripts were found in U2OS cells overexpressing FRG1 as well as transgenic mice overexpressing FRG1, implicating FRG1 is involved in alternative splicing (Gabellini et al., 2006; van Koningsbruggen et al., 2007). Biochemically, FRG1 has been found to associate with RNA processing components and large-scale proteomic studies revealed that FRG1 is associated with human spliceosome complex (Rappsilber et al., 2002; van Koningsbruggen et al., 2007). Interestingly, two potential FRG1-associating proteins, which are RNA processing components, SMN and PABPN1, are themselves involved in neuromuscular disorders (Brais et al., 1998; Briese et al., 2005; Calado et al., 2000).

FRG1, due to its location 125 kb centromeric to the genetic lesion of FSHD, the D4Z4 repeats, is one of the early candidate genes (van Deutekom et al., 1996). However, expression levels of FRG1 in muscle biopsies, primary myoblasts and myotubes derived from FSHD patients have varied from 25-fold increased, unchanged, and 5-fold decreased compared with controls (Arashiro et al., 2009; Gabellini et al., 2002; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; Winokur et al., 2003). Any role for FRG1 in FSHD pathogenesis is further complicated by our lack of understanding of the normal function of FRG1 in muscle development.

We have previously shown that FRG1 is a dynamic nuclear and cytoplasmic protein in Chapter 2. This chapter will further investigate the cellular localization of endogenous human FRG1 and focus on the nuclear aspects of human FRG1 biology in respect to RNA biogenesis. We further characterized the FRG1 nuclear and nucleolar localizations and identified FRG1 RNA association *in vitro* and *in vivo*, and show that

FRG1 is a dynamic nuclear-cytoplasmic shuttling protein that interacts with TAP, a nuclear RNA export protein.

RESULTS

Endogenous human FRG1 is both nuclear and cytoplasmic, and is prominently localized to the granular component of nucleoli

In order to characterize the endogenous human FRG1, HeLa cells were subjected to immunocytochemistry using the highly specific HS1 FRG1 antibody described in Chapter 2 (Figure 3.1). In contrast to the exclusively nuclear localization often seen for overexpressed, epitope-tagged FRG1 (van Koningsbruggen et al., 2004), the endogenous HeLa FRG1 exists in both nuclear and cytoplasmic pools. Consistent with previous data for the overexpressed FRG1, the endogenous nuclear FRG1 is prominently localized in the DAPI-poor foci (Figure 3.1A-C). Co-immunostaining for FRG1 and fibrillarin, a nucleolar marker (Aris and Blobel, 1991) showed the endogenous FRG1 concentrates in the nucleoli (Figure 3.1D-G), however, the nucleolar FRG1 associated with but did not completely co-localize with fibrillarin (Figure 3.1D-G) suggesting that nucleolar FRG1 was not part of the dense fibrillar component (DFC). The DFC and granular center (GC) of the nucleolus can be readily distinguished from each other by inhibition of RNA polymerase I transcription using Actinomycin D (AMD) (Olson and Dundr, 2005). HeLa cells treated with AMD showed FRG1 localized to the GC, distinctly separate from the DFC marked by fibrillarin (Figure 3.1H-K). Beyond the nucleolar localization, FRG1 showed punctate nuclear immunostaining suggestive of nuclear speckles (Figure 3.1E, I, M), therefore, HeLa cells were co-immunostained for FRG1 and the nuclear speckle

marker SC35 (Spector et al., 1991). FRG1 did not co-localize with SC35 indicating it is not a part of SC35-containing nuclear speckles (Figure 3.1L-O).

Endogenous FRG1 associates with nascent RNA transcripts in vivo

FRG1 was previously implicated in aspects of RNA biogenesis, in particular pre-mRNA splicing. It is currently well accepted that most RNA processing factors are recruited to nascent transcripts co-transcriptionally, resulting in the formation of export-competent ribonucleoprotein (RNP) complexes at the sites of transcription. Many of these factors, such as the heterogenous nuclear ribonucleoprotein G (hnRNP G) were shown to associate with most transcription units of RNA polymerase II (RNAPII) in amphibian oocytes (Morgan, 2002; Soulard et al., 1993). Chromatin in *Xenopus laevis* oocytes is structured within lampbrush chromosomes (LBC) and several thousand non-chromosomal nucleoli. LBCs are often described as extended diplotene bivalent chromosomes, and their characteristic shape is the result of the intense transcriptional activity that is associated with them. Each homologue consists of a condensed chromatin axis, from which are escaping numerous pairs of lateral loops that are active transcription units of RNAPII. Although the chromatin of these loops almost certainly adopts the conformation of a 10nm (or less) fiber (Morgan, 2002), they are readily distinguishable by light microscopy because the chromatin axis is surrounded by a dense matrix of nascent RNP fibrils. We took advantage of this unique spatial resolution to test whether FRG1 associates with nascent transcripts, using an antibody specific for *X. laevis* FRG1 (Hanel et al., 2009) on nuclear spreads (Figure 3.2). Endogenous FRG1 is detected within the RNP matrix of the LBC loops (Figure 3.2E-L), which provides strong evidence that FRG1

interacts with nascent transcripts within the cell nucleus. In addition, FRG1 is also found associated with Cajal bodies, nucleoli and the nuclear speckles (or interchromatin granule clusters) (Figure 3.2A-D), which is consistent with the subnuclear distribution of FRG1 defined previously in human cells and further support a role of FRG1 in RNA biogenesis (van Koningsbruggen et al., 2004).

To determine if the endogenous human FRG1 associates with human mRNAs *in vivo*, RNA immunoprecipitations (RIP) were performed on HeLa whole cell extracts using the HS1 FRG1-specific antibody previously characterized (M. L. Hanel, submitted for publication). After the RIP procedure, FRG1 associated RNAs were subjected to reverse transcriptase PCR (RT-PCR), initially assaying for FXR1, a transcript known to be affected in FSHD (Davidovic et al., 2008) (Figure 3.3, lanes 2 and 6). All RIP RT-PCR products were sensitive to RNase A treatment of the starting material confirming the RNA (and not DNA) association (Figure 3.3, lanes 7 and 9). Furthermore, a RIP non-specific antibody control, affinity-purified PAT-9 antiserum was used and failed to IP any detectable RNAs illustrating the specificity of the RIP procedure for FRG1 (Figure 3.3, lane 3). Subsequently, RIP reactions were assayed by RT-PCR for the FRG1 and HPRT mRNAs, all producing results similar to that shown for FRX1 (data not shown). Although it is not clear if the FRG1-RNA association is direct or indirect, these results show that endogenous human FRG1 associates with numerous mRNA transcripts *in vivo*.

Recombinant human FRG1 interacts with RNA directly

Although FRG1 does not contain any conventional RNA-binding motif, many RNA associated ribonucleoproteins (RNPs) interact with RNA directly *in vitro* (Dreyfuss

et al., 2002; Dreyfuss et al., 1993). To determine if FRG1 is capable of binding with RNA directly, RNA electrophoretic mobility shift assays (REMSA) were performed using recombinant human FRG1 or deletion mutants of FRG1 (Figure 3.4A and B) and radiolabeled, *in vitro* transcribed *Xenopus* β-globin RNA as a probe. Full-length FRG1 formed a protein-RNA complex as indicated by the FRG1-dependent slower migrating RNA species (Figure 3.4C, lanes 2-4). Since there is no predicted RNA binding domain described for FRG1, recombinant proteins from four deletion constructs were used for REMSA to identify the FRG1 RNA-binding domain. The results showed that FRG1 binds RNA directly through its N-terminal NLS as deleting this region abolishes the RNA-binding activity of FRG1 (Figure 3.4C, lanes 14-16) while the other three deletions maintained RNA-binding activity similar to full-length FRG1. To determine if this interaction was specific to RNA, competition assays were performed using *in vitro* transcribed but unlabeled RNA probe as a specific competitor, and tRNA and double stranded (ds) DNA as query competitors. The FRG1-RNA interaction was successfully competed by the addition of cold RNA probe (Figure 3.4D, lanes 6-9) and tRNA (Figure 3.4D lanes 10-13), however, dsDNA did not compete (Figure 3.4D, lanes 14-17) indicating that the FRG1-nucleic acid interaction is specific to RNA.

FRG1 shuttling is dependent on RNA polymerase II-mediated transcription

FRG1 has been shown to shuttle between the nucleus and cytoplasm in Chapter 2, and to associate with mRNAs (Figures 3.2-4), therefore the effect of transcription inhibition on FRG1's nuclear shuttling was investigated by heterokaryon assay (Figure 3.5). Transfected HeLa cells, readily identified by Hoechst staining as containing large

DNA-poor nucleoli (Figure 3.5, blue arrows), that transiently express HA-FRG1 were fused to non-transfected C2C12 cells, easily identified by the multiple DNA-bright foci (Figure 3.5, white arrows), in presence of cycloheximide (CHX) blocking *de novo* translation. Overexpressed HA-FRG1 in the HeLa cells showed exclusive nuclear localization with enrichment at nucleoli as reported previously (van Koningsbruggen et al., 2004). Three hours after fusion was initiated, HA-FRG1 staining was detected in C2C12 nuclei (Figure 3.5A-D), and this shuttling was not affected by Leptomycin B (LMB), an inhibitor of the CRM1/exportin1 nuclear export pathway (Figure 3.5E-H compared with A-D, CHX alone). These results indicate FRG1 nuclear export occurs through a CRM1 independent pathway. We further investigated an effect of the transcription on FRG1 shuttling. Treatment with low levels of AMD (0.4 mg/ml) had no apparent effect on FRG1 shuttling or nucleolar localization in HeLa cells (Figure 3.5I-L); however, treatment with a higher dose of AMD (5 μg/ml), inhibiting all three RNA polymerases, resulted in FRG1 being excluded from HeLa nucleoli and appeared consistently to be preferentially accumulated into the nuclei of recipient C2C12 cells (Figure 3.5M-P). These data suggest FRG1 nuclear export is somehow linked to RNA polymerase II-mediated transcription.

FRG1 interacts with the nuclear export protein TAP

Considering FRG1 binds RNA *in vitro,* associates with mRNA *in vivo*, and its nuclear shuttling depends on RNA polymerase II transcription, it seemed reasonable that FRG1 might be involved in mRNA export as part of a hnRNP complex. Since the majority of mature mRNAs in the nucleus are exported through the TAP/NXT1 pathway,

FRG1 was tested for a direct interaction with TAP *in vitro* (Figure 3.6). GST pull-down assays were performed using recombinant GST-TAP and radiolabeled *in vitro* transcribed and translated (TnT) full-length human FRG1 and FRG1 deletion constructs. The results showed that FRG1 interacts directly with TAP mediated through the FRG1 N-terminal NLSs (Figure 3.6A). Since both TAP and FRG1 interact with RNA, and RNA could be present from the recombinant protein purification or TnT reaction, the GST pull-down assay was repeated with benzonase nuclease to digest any RNA or DNA. Benzonase nuclease treatment did not affect the FRG1-TAP interaction indicating that FRG1 interacts with TAP directly and independent of RNA (Figure 3.6B).

DISCUSSION

FRG1 is critical for development of the vertebrate musculature and vasculature and has been implicated in mediating FSHD pathophysiology. Still, very little is known about FRG1, hindering our understanding of how changes in its expression levels might lead to disease. In this Chapter, we further investigated human FRG1 by characterizing its subcellular localizations, biological activities and identifying new interacting proteins.

Experiments were performed to determine the biological function of nuclear FRG1. The nuclear fraction of endogenous FRG1 is concentrated in Cajal bodies (CBs) and the granular component (GC) of nucleoli, as well as associated with nascent mRNA transcripts and the actively transcribed regions of the euchromatin. The CBs are nuclear bodies rich in factors involved in transcription and RNA processing and are sites of snRNP biogenesis (Austin and Bellini, 2010; Gall, 2003). Thus, FRG1's presence in CBs is consistent with its proposed role in RNA biogenesis and provides additional similarity

with RNPs (Rappsilber et al., 2002; van Koningsbruggen et al., 2004; van Koningsbruggen et al., 2007). Similarly, FRG1's nucleolar localization supports these functional roles. Nucleoli, the dynamic structures containing tandem chromosomal repeats encoding rRNAs, are sites where pre-ribosomal particles are transcribed and assembled; however, nucleoli have additional roles including maturation of some RNP particles, assembly of the RNA splicing machinery, and sequestration of certain nuclear regulatory factors (Austin and Bellini, 2010; Boisvert et al., 2007). FRG1's nucleolar localization to the GC and not with either fibrillar component suggests FRG1 is involved in later stages of rRNA processing and not directly with rDNA transcription (Sirri et al., 2008). In addition to the rRNA co-localization, both lampbrush chromosome spreads and RNA-IPs from HeLa cells indicate that the endogenous FRG1 associates with nascent RNA polymerase II generated mRNAs *in vivo*. However, although FRG1 is capable of directly interacting with RNA *in vitro*, and is in a complex with mRNA *in vivo*, it is not clear that *in vivo* its RNA association is mediated directly through FRG1. Interestingly in FSHD-derived muscle cells, the muscle-specific isoforms of FXR1, a target of FRG1, show an aberrant expression pattern due to decreased mRNA stability (Davidovic et al., 2008) while the mRNAs from fast skeletal muscle troponin T (*TNNT3*) and myotubularin related protein 1 (*MTMR1*) show aberrant alternative mRNA splicing (Gabellini et al., 2006).

Many RNA-binding proteins are involved in pre-mRNA processing as well as the transport, localization, translation and the stability of mRNAs (Dreyfuss et al., 2002). Based on our data showing the endogenous FRG1 is both nuclear and cytoplasmic, associates with mRNAs, and is a nuclear shuttling protein, we thus hypothesized that

FRG1 might function in the nuclear to cytoplasmic transport of mRNAs. We found that FRG1 interacts directly with TAP/NXF1, a member of the NXF family of transport receptors (Gruter et al., 1998). TAP-NXT1 heterodimers associate with RNA-binding adaptor proteins in the nucleus to mediate the nuclear export of the majority of mature mRNAs through the nuclear pore complex (NPC). Together with previous data identifying FRG1 as a component of the spliceosome (Rappsilber et al., 2002), these data support a role for FRG1 in several aspects of RNA biogenesis, including RNA splicing and RNA transport.

MATERIALS AND METHODS

DNA constructs

All PCR primers are listed in Table 3.1. To generate the FRG1 His-tagged bacterial expression constructs, the cDNA for the full-length human FRG1 coding sequence was PCR amplified (primers #1 and #2) and subcloned between the NdeI and XhoI restriction sites into pET-23b vector (Novagen, Gibbstown, NJ). The pET-FRG1 deletion constructs were made by the same procedure with the following primer sets: #2 and #3 for (Δ1-32), #2 and #4 for (Δ1-20), #1 and #5 for (Δ235-258), and #1 and #6 for (Δ183-258).

Cell culture

HeLa, and C2C12 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 2 mM L-glutamine, supplemented with 10% fetal bovine serum

(FBS) and 1% penicillin–streptomycin (pen/strep). Cells were incubated with 5% CO_2 at 37°C.

Antibodies

Four rabbit polyclonal antibodies were used, all raised against synthetic peptides corresponding to human FRG1. The HS1 (NH_2-CKKDDIPEEDKGNVK) and HS2 (NH_2-CGRSDAIGPREQWEP) FRG1 antibodies were generated by GenScript USA Inc. (Piscataway, NJ), and antisera were affinity purified against the peptide cross-linked to NHS-Sepharose (GE Healthcare, Piscataway, NJ), eluted in 10 mM glycine, pH 2.5, dialyzed against PBS pH7.4, and characterized in Chapter 2.

Immunocytochemistry

HeLa cells were fixed in 2% formaldehyde (FA) in PBS for 15 min at room temperature (RT). After fixation, cells were made permeable with 0.25% Triton-X 100 in PBS for 10 min on ice, and subsequently blocked with 2% BSA in PBS for 30 min at RT. Primary antibody incubation was carried out overnight at 4°C in PBS with 2% BSA, and secondary antibody incubation was for 40 min at RT in PBS. Co-staining experiments were performed using FRG1 HS1, mouse monoclonal fibrillarin (ab18380, Abcam, Cambridge, MA) and SC-35 (S4045, Sigma-Aldrich, St. Louis, MO) antibodies at 1:500, 1:100 and 1:200 dilutions, respectively. Secondary antibodies used were FITC-conjugated donkey anti-rabbit (pre-cleared) and rhodamine-conjugated goat anti-mouse, or rhodamine-conjugated goat anti-rabbit and FITC-conjugated goat anti-mouse (Jackson

ImmunoResearch Laboratories Inc) used at 1:100. The 4,6-diamidino-2-phenylindole (DAPI) was used at 1µg/ml.

Lampbrush chromosome spreads

Female adult frogs (*Xenopus laevis*) were anesthetized in 0.15% tricaine methanesulfonate (Sigma-Aldrich), and small fragments of ovary were surgically removed. Oocytes were defolliculated for 1 hr in saline buffer OR2 (Wallace et al., 1973) containing 0.15% collagenase type II (Sigma-Aldrich). Stage IV-V oocytes were selected and maintained in OR2 at 18˚C. Nuclear spreads were prepared as described (Patel et al., 2007). Fixed nuclear spreads were rinsed in PBS and blocked in PBS containing 0.5% bovine serum albumin (Sigma-Aldrich) plus 0.5% gelatin (from cold-water fish). Spreads were incubated with primary antibody, XTB-FRG1 (Hanel et al., 2009), for 1 hr at RT, washed for 30 min with two changes of PBS, incubated with secondary antibody Alexa 488 conjugated goat anti-rabbit antibody (Invitrogen Corp.) for 1 hr at RT, and washed again for 30 min with two changes of PBS. Spreads were mounted in 50% glycerol containing 1mg/ml phenylenediamine and 10 pg/ml DAPI.

Nuclear shuttling assay

The assay was carried out essentially as described (M. L. Hanel, submitted for publication). HeLa cells (~60% confluent) were transfected with pcDNA3.1HA-FRG1 (M. L. Hanel, submitted for publication) using TransIT-LT1 reagent (Mirus Bio, Madison, WI) and allowed to grow for 24 hrs. The cells were removed with trypsin, washed in PBS, plated (1×10^6) on glass coverslips and allowed to adhere for 2 hrs before non-

transfected murine C2C12 cells (5 x 10^5) were overlaid onto the transfected HeLa cells for 3 hours. Translation was stopped with cycloheximide (CHX) (100 μg/ml) treatment for 15 min prior to fusion. After removing the media, cells fused by adding 50% PEG 4000 in DMEM for 2 min. The fusions were immediately washed with DMEM and then incubated for three hours in the presence of CHX (100 μg/ml), CHX (100 μg/ml) + actinomycin D (AMD) (0.4 μg/ml low; 5.0 μg/ml high), or CHX (100μg/ml) + leptomycin B (2 ng/ml) and processed for immunocytochemistry using a HA (clone 3F10, Roche) monoclonal antibody and co-staining with Hoechst 33342 (5 μg/ml).

RNA-IP

These experiments were carried out essentially as described (Lin et al., 2005). HeLa cells (1x10^8) were suspended in 40 ml 1x PBS with 4 ml cross-linking buffer (100 mM NaCl, 50 mM HEPES, 1 mM EDTA, 0.5 mM EGTA, 11% formaldehyde) for 30 min at RT with rocking. The reaction was quenched with 2.2 ml 2.5 M glycine (pH 7.0) for 5 min at RT then washed with 1X PBS and pelleted. Cells were lysed by suspending in FA buffer (50 mM HEPES-KOH [pH 7.5], 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, protease inhibitors) + RNase inhibitor (50 U/500 ml) and sonicated on a Branson sonifier at 25% power for 20 pulses of 20 sec with 30 rest on ice between each pulse. The solution was made up to 25 mM $MgCl_2$ and 5 mM $CaCl_2$ then RNase-free DNase I (100 U/500ml) and RNase inhibitor (3 ml/500 ml) were added and incubated at 37°C for 30 min. The reactions were centrifuged at 20,000 *g* for 15 min at 4°C, keeping the supernatant for IPs.

For each RNA-IP reaction, 100 µl cleared extract was diluted with 900 ml ChIP buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, and 0.5% NP-40) to which 6 ml antibody (HS1-FRG1 or aPAT-9) and 4 ml RNase inhibitor were added then rotated for 12 hrs at 4°C. Protein A Dynabeads (40 ml/IP) were washed twice in ChIP buffer then added to the IP reactions for 1 hr. Dynabead IPs were washed 3X in 1 ml wash buffer (50 mM Tris pH 7.4, 500 mM NaCl, 1% Triton X-100, 0.1% SDS) and RNA was eluted in 200 ml Elution Buffer (200 mM NaCl, 50 mM Tris pH 7.4, 20 mg Proteinase K) for 1 hr at 42°C. Cross-linking was reversed by heating at 65°C for 5 hrs, and the RNA was extracted with acid equilibrated (pH 4.8) phenol:chloroform (5:1), ethanol precipitated, and brought up in DEPC-dH_2O for use in RT-PCR (primers: #7, FXR-F and #8, FXR-R, Table 3.1).

RNA electrophoretic mobility shift assay (REMSA)

To generate full-length and deletion-containing FRG1 recombinant proteins, the pET23 plasmid constructs described above were transformed into E. coli BL21(DE3) cells and induced with 1 mM IPTG. The protein was purified using TALON resin (Clontech, Mountain View, CA) as per the manufacturer's instructions. The T7TS plasmid, containing the 5' and 3' untranslated regions (UTRs) of the *Xenopus* β-globin mRNA, was used as template for making mRNA probes. T7TS was linearized with Bgl II (New England Biolabs, NEB, Ipswich, MA) and *in vitro* transcribed using T7 RNA polymerase (NEB) and [^{32}P]-UTP according to the manufacturer's instructions, treated with TURBO DNase (Ambion, Inc., Austin, TX) for 15 min at 37˚C, purified on a 5% TBE gel and the extracted radiolabeled RNA was quantified by scintillation counter (Beckman LS6500).

The RNA-protein binding assay was performed in REMSA buffer (1 mM $MgCl_2$, 10 mM Tris 7.4, 5% glycerol, 0.1% IGEPAL/NP-40, 50 mM NaCl, and 20 u RNasin (Promega, Madison, WI) with 50,000 CPM [^{32}P] labeled probe. Recombinant full-length FRG1 and deletion constructs were used at 150, 300, and 450 ng. Reactions were incubated on ice for 20 min, analyzed by 5% TBE gel, and visualized by autoradiography. Unlabeled RNA probe, tRNA (Sigma-Aldrich), and T7TS plasmid, added to the reactions prior to the specific radiolabeled probe, were used as competitors.

GST pull-down

To test the FRG1-TAP interaction, recombinant GST and GST-Tap proteins were purified and bound to glutathione beads as described (Matzat et al., 2008). Meanwhile, pET23-FRG1 constructs were used for *in vitro* transcription/translation (TnT) according to the manufacturer's instructions (TnT-kit, Promega). After washing the glutathione beads bound to GST proteins, the ^{35}S methionine labeled proteins were added and subsequently incubated with the GST or GST-TAP beads with rotation for 3 hrs at 4°C. The FRG1-TAP interactions were carried out in 150 mM NaCl as described (Matzat et al., 2008). The beads were washed and proteins eluted by boiling with 20 µl of 2x SDS-sample buffer, after which the whole samples were loaded onto a SDS-page gel. The gels were dried and exposed to film overnight.

Microscopy

Standard fluorescence microscopy was carried out using a HCL FL Fluotar 100X oil objective (NA=1.30) on an upright Leica DMR microscope. Images were captured

using a monochrome Retiga EXI Charge-Coupled Device (CCD) camera (Qimaging) driven by the In vivo software (version 3.2.0, Media Cybernetics). All images were processed using Adobe Photoshop to adjust brightness, contrast, size, and merged or split channels. Applied Precision Personal Deltavision was used for deconvolution images.

REFERENCES

Arashiro, P., I. Eisenberg, et al. 2009. Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. Proc Natl Acad Sci U S A, **106**: 6220-5.

Aris, J. P. and G. Blobel. 1991. cDNA cloning and sequencing of human fibrillarin, a conserved nucleolar protein recognized by autoimmune antisera. Proc Natl Acad Sci U S A, **88**: 931-5.

Austin, C. M. and M. Bellini. 2010. The dynamic landscape of the cell nucleus. Mol Reprod Dev, **77**: 19-28.

Boisvert, F. M., S. van Koningsbruggen, et al. 2007. The multifunctional nucleolus. Nat Rev Mol Cell Biol, **8**: 574-85.

Brais, B., J. P. Bouchard, et al. 1998. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. Nat Genet, **18**: 164-7.

Briese, M., B. Esmaeili, et al. 2005. Is spinal muscular atrophy the result of defects in motor neuron processes? Bioessays, **27**: 946-57.

Calado, A., F. M. Tome, et al. 2000. Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. Hum Mol Genet, **9**: 2321-8.

Davidovic, L., S. Sacconi, et al. 2008. Alteration of expression of muscle specific isoforms of the fragile X related protein 1 (FXR1P) in facioscapulohumeral muscular dystrophy patients. J Med Genet, **45**: 679-85.

Dreyfuss, G., V. N. Kim, et al. 2002. Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol, **3**: 195-205.

Dreyfuss, G., M. J. Matunis, et al. 1993. hnRNP proteins and the biogenesis of mRNA. Annu Rev Biochem, **62**: 289-321.

Gabellini, D., G. D'Antona, et al. 2006. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. Nature, **439**: 973-7.

Gabellini, D., M. R. Green, et al. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell, **110**: 339-48.

Gall, J. G. 2003. [A role for Cajal bodies in assembly of the nuclear transcription machinery]. Tsitologiia, **45**: 971-5.

Grewal, P. K., L. C. Todd, et al. 1998. FRG1, a gene in the FSH muscular dystrophy region on human chromosome 4q35, is highly conserved in vertebrates and invertebrates. Gene, **216**: 13-9.

Gruter, P., C. Tabernero, et al. 1998. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. Mol Cell, **1**: 649-59.

Hanel, M. L., R. D. Wuebbles, et al. 2009. Muscular dystrophy candidate gene FRG1 is critical for muscle development. Dev Dyn, **238**: 1502-12.

Jiang, G., F. Yang, et al. 2003. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet, **12**: 2909-21.

Klooster, R., K. Straasheijm, et al. 2009. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. Eur J Hum Genet, **17**: 1615-24.

Kureishy, N., V. Sapountzi, et al. 2002. Fascins, and their roles in cell structure and function. Bioessays, **24**: 350-61.

Lin, C., L. Yang, et al. 2005. ATPase/helicase activities of p68 RNA helicase are required for pre-mRNA splicing but not for assembly of the spliceosome. Mol Cell Biol, **25**: 7484-93.

Liu, Q., T. I. Jones, et al. 2010. Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. J Cell Sci, **123**: 1116-23.

Masny, P. S., O. Y. Chan, et al. 2010. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. Eur J Hum Genet, **18**: 448-56.

Matzat, L. H., S. Berberoglu, et al. 2008. Formation of a Tap/NXF1 homotypic complex is mediated through the amino-terminal domain of Tap and enhances interaction with nucleoporins. Mol Biol Cell, **19**: 327-38.

Morgan, G. T. 2002. Lampbrush chromosomes and associated bodies: new insights into principles of nuclear structure and function. Chromosome Res, **10**: 177-200.

Olson, M. O. and M. Dundr. 2005. The moving parts of the nucleolus. Histochem Cell Biol, **123**: 203-16.

Osborne, R. J., S. Welle, et al. 2007. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology, **68**: 569-77.

Patel, S. B., N. Novikova, et al. 2007. Splicing-independent recruitment of spliceosomal small nuclear RNPs to nascent RNA polymerase II transcripts. J Cell Biol, **178**: 937-49.

Rappsilber, J., U. Ryder, et al. 2002. Large-scale proteomic analysis of the human spliceosome. Genome Res, **12**: 1231-45.

Sirri, V., S. Urcuqui-Inchima, et al. 2008. Nucleolus: the fascinating nuclear body. Histochem Cell Biol, **129**: 13-31.

Soulard, M., V. Della Valle, et al. 1993. hnRNP G: sequence and characterization of a glycosylated RNA-binding protein. Nucleic Acids Res, **21**: 4210-7.

Spector, D. L., X. D. Fu, et al. 1991. Associations between distinct pre-mRNA splicing components and the cell nucleus. EMBO J, **10**: 3467-81.

van Deutekom, J. C., R. J. Lemmers, et al. 1996. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. Hum Mol Genet, **5**: 581-90.

van Koningsbruggen, S., R. W. Dirks, et al. 2004. FRG1P is localised in the nucleolus, Cajal bodies, and speckles. J Med Genet, **41**: e46.

van Koningsbruggen, S., K. R. Straasheijm, et al. 2007. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. Chromosoma, **116**: 53-64.

Wallace, R. A., T. Ho, et al. 1973. Protein incorporation by isolated amphibian oocytes. IV. The role of follicle cells and calcium during protein uptake. Exp Cell Res, **82**: 287-95.

Winokur, S. T., K. Barrett, et al. 2003. Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. Neuromuscul Disord, **13**: 322-33.

Wuebbles, R. D., M. L. Hanel, et al. 2009. FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. Dis Model Mech, **2**: 267-74.

FIGURES AND TABLES

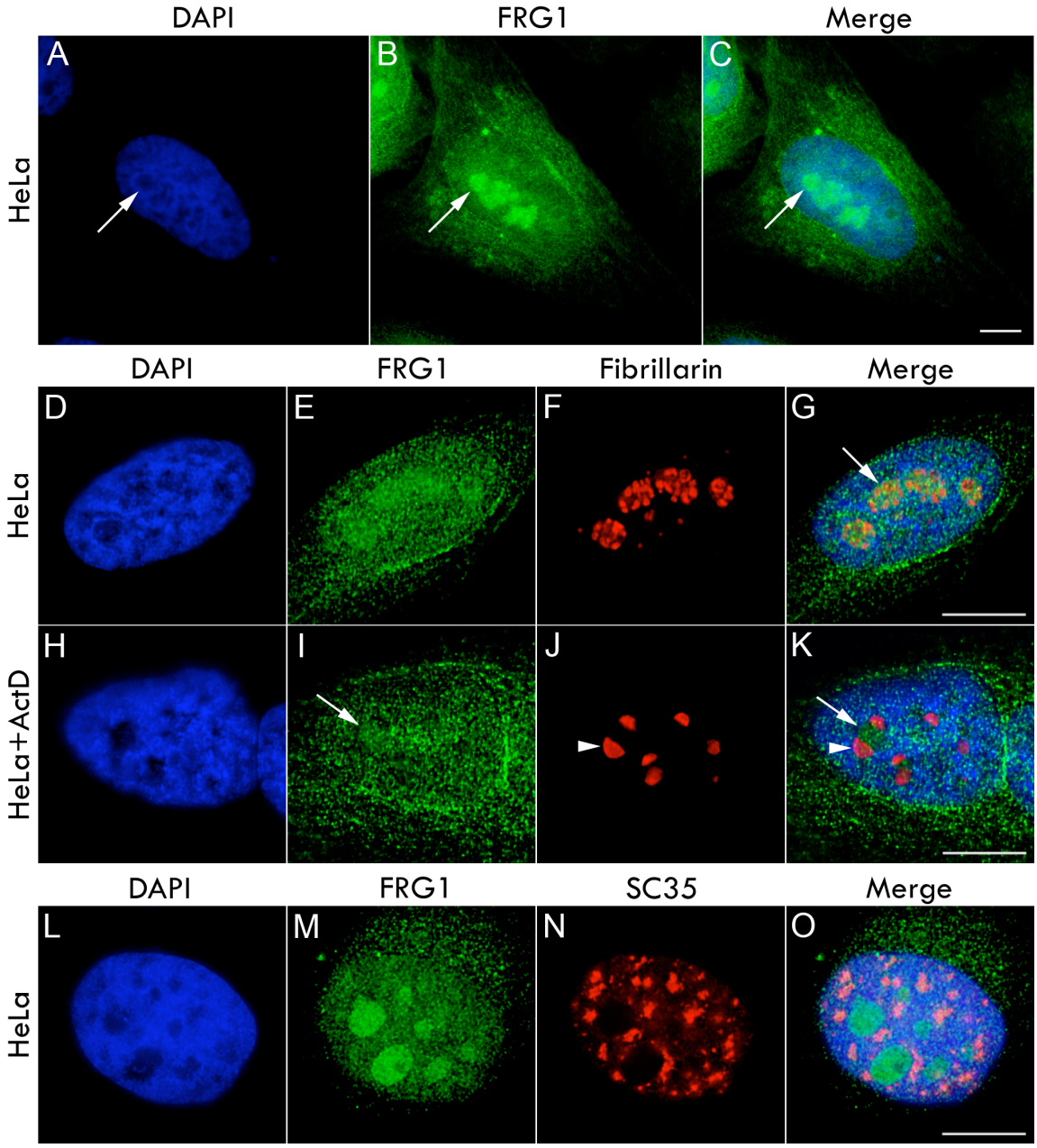

Figure 3.1: Endogenous human FRG1 is both nuclear and cytoplasmic in HeLa cells, and nuclear FRG1 localizes predominantly to the nucleoli but not nuclear speckles. (A-C) Immunostaining using the HS1 FRG1 antibody (green) shows the endogenous FRG1 is both nuclear and cytoplasmic. Nuclear FRG1 is prominent in the DAPI-poor nucleoli (arrows). (D-G) Co-immunostaining for FRG1 (green) and the nucleolar marker fibrillarin (red) confirms the nucleolar localization of FRG1. (H-K) HeLa cells treated with ActD show FRG1 localizes to the granular component (GC) of the nucleoli (arrow) and not the fibrillarin stained dense fibrillar component (DFC) (arrowhead). (L-O) HeLa cells co-immunostained with FRG1 (green) does not co-localize with nuclear speckles indicated by SC35 staining (red). Bars = 10 μm.

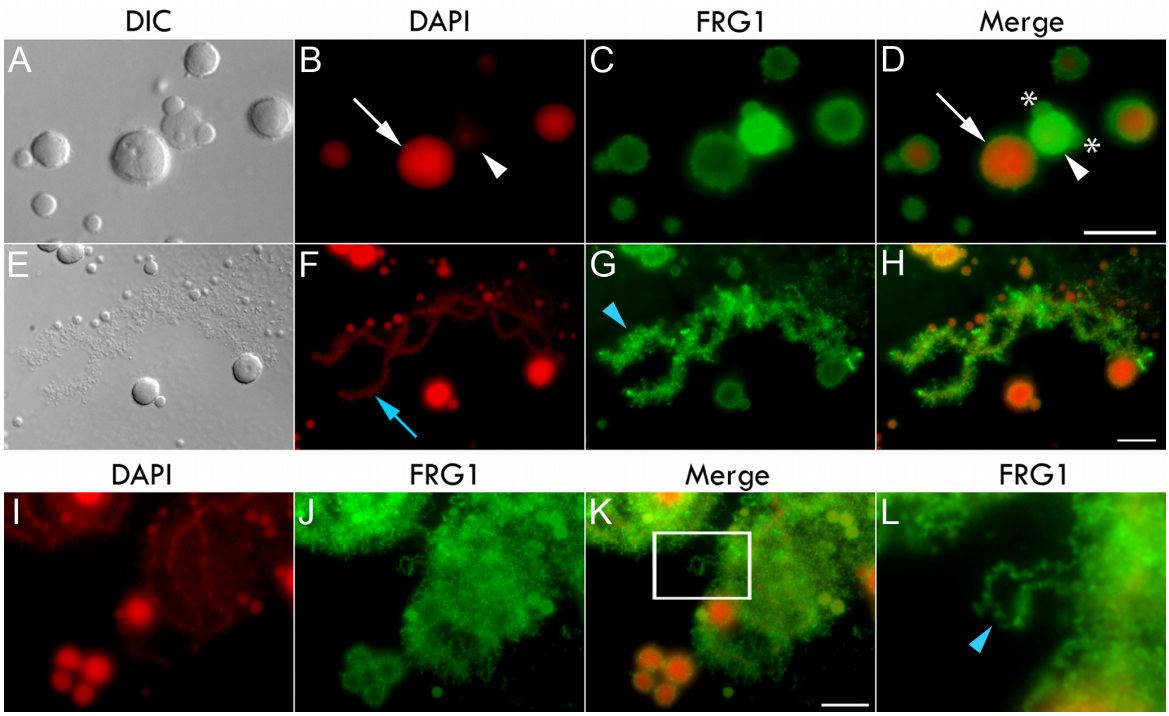

Figure 3.2: Endogenous FRG1 associates with nucleoli, Cajal bodies, and nascent RNA transcripts on lampbrush chromosome spreads in *Xenopus laevis* oocytes. *X. laevis* oocyte nuclei isolated under oil and immunostained using antibodies specific for *X. laevis* FRG1 show FRG1 (green) localizes to (A-D) the nucleoli (arrow), Cajal bodies (arrowhead) with associating B snurposomes (indicated by *), and (E-L) nascent mRNAs that are being actively transcribed (blue arrowhead). DAPI staining (red) labels the DNA, marking the chromosome axis (blue arrow in F) and distinguishing nucleoli from Cajal bodies (B). (I-L) Focus on FRG1 localizing to the transcriptionally active chromosome loops, indicated by the blue arrowhead in L. The boxed area in K is shown magnified 3X in L. Bars = 10 µm.

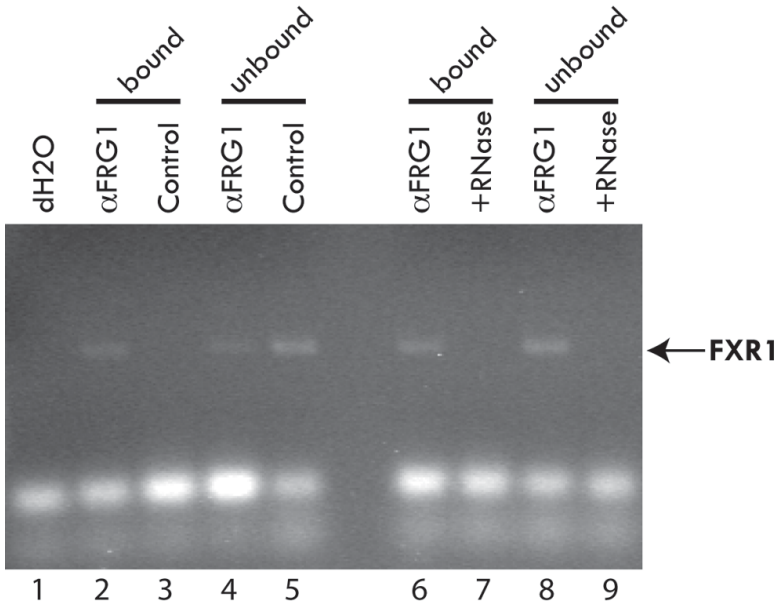

Figure 3.3: Endogenous human FRG1 is associated with mRNA *in vivo*. RNA-IPs using the HS1 FRG1 antibody show FRG1 associates with FXR1 mRNA transcripts *in vivo* in HeLa cells as assayed by RT-PCR. The FRG1 antibody specifically IP'ed FXR1 mRNA (lanes 2 and 6) while a control non-specific (αPAT-9) antibody (lanes 3) did not. The unbound fraction confirms the presence of FRX1 mRNA in the control reaction (lane 5). The RT-PCRs are sensitive to RNase treatment (lanes 7 and 9), confirming that RNA is the IP'ed nucleic acid.

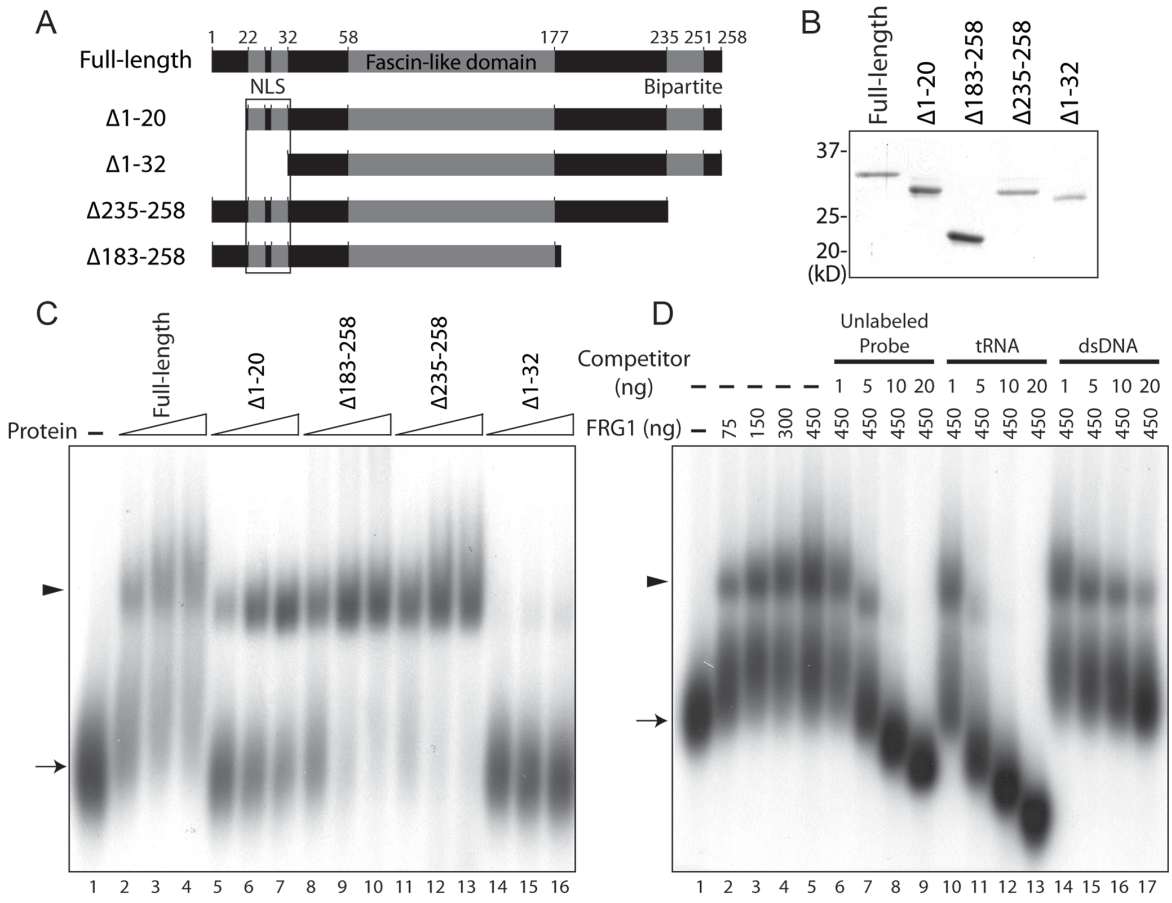

Figure 3.4: Human FRG1 interacts with RNA directly through its N-terminal NLS. (A) The schematic diagram showing protein domains of full-length human FRG1 and the deletion constructs. (B) The recombinant proteins of full-length human FRG1 and deletions shown by Coomassie blue staining of SDS-PAGE gel. (C) RNA gel shift experiment using radiolabeled, *in vitro* transcribed mRNA as probe (arrow) showed FRG1 forms Protein-RNA complexes (arrowhead) through its N-terminal NLS as deleting this region abolished FRG1-RNA interaction. (D) The FRG1-RNA interaction is specific as both unlabeled probe and tRNA compete with radiolabeled probe while dsDNA does not. Box in A shows the region responsible for FRG1-RNA interaction.

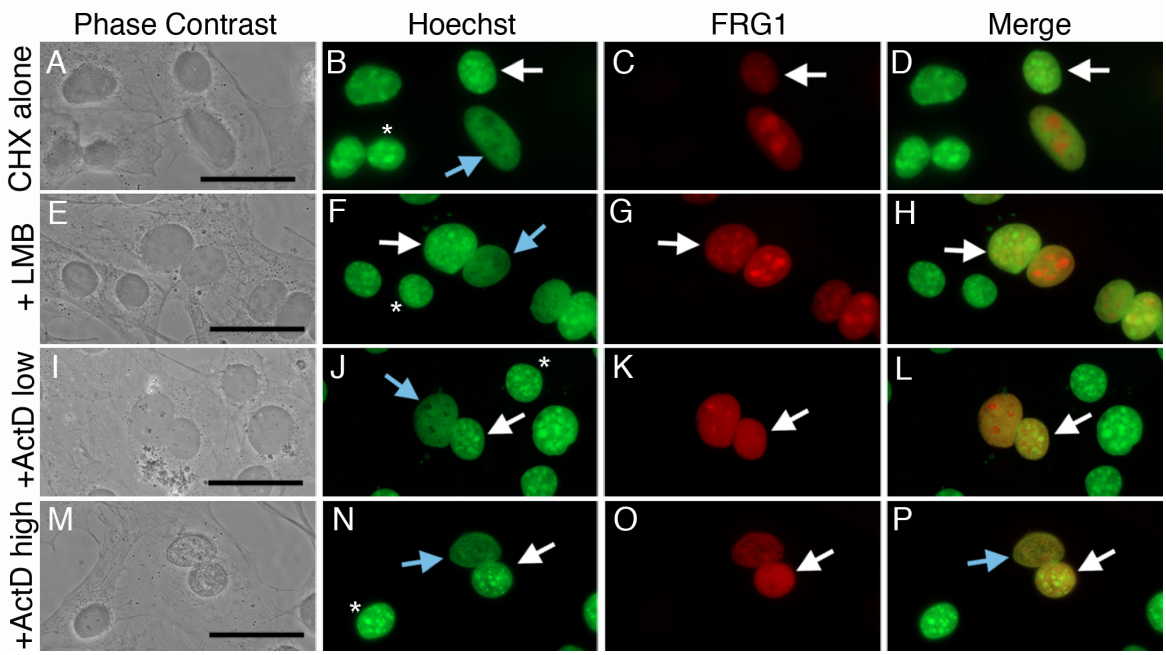

Figure 3.5: FRG1 nuclear-cytoplasmic shuttling is independent of LMB but affected by inhibition of RNA polymerase II-mediated transcription. HeLa cells, distinguished by clear DNA-poor regions (blue arrow), expressing HA-FRG1 were fused to murine C2C12 cells distinguished by DNA-rich foci (white arrow) in the presence of the translation inhibitor CHX for 3 hours with or without additional treatments described. Unfused C2C12 cells (examples indicated by *) do not express HA-FRG1. (A-D) HA-FRG1 expressed in HeLa cells localizes to C2C12 nuclei 3 hours after inducing fusion, showing that FRG1 shuttles between the nucleus and the cytoplasm. Additional cell fusion experiments were carried out in the presence of LMB (E-H), low levels of AMD (I-L) to shut down RNA polymerase I mediated transcription, and high levels (M-P) of AMD to shut down RNA polymerase I and II mediated transcription, respectively. LMB or low does of AMD treatment do not affect FRG1 shuttling. Under high AMD concentration, HA-FRG1 accumulated in the recipient C2C12 cells, and was excluded from the HeLa nucleoli. Bars = 10 μm.

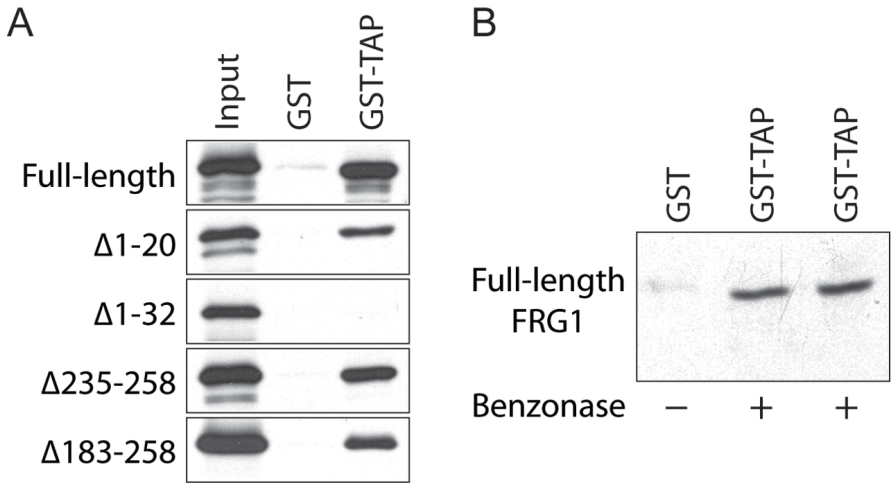

Figure 3.6: FRG1 directly associates with TAP through the N-terminal NLS. (A) Radio-labeled, *in vitro* transcribed and translated FRG1 and deletion constructs of FRG1 were incubated with GST-TAP or GST alone. Bound proteins were resolved by SDS-PAGE followed by autoradiography. (B) The FRG1-TAP interaction is independent of nucleic acids as the interaction is insensitive to benzonase nuclease treatment.

Table 3.1 PCR primers

Primers used for making RBD deletions and RT-PCR

#1: CCAACATATGGCCGAGTACTCCTACGTGAAGTC

#2: CCAACTCGAGCTTGCAGTATCTGTCGGCTTTC

#3: CATATGAGAGAAGAAGATGAAGAAACCCAGCTTG

#4: CATATGAGTAAGAAGAAAAAGAGCAAAGATAAGAAA

#5: CTCGAGAGCCTTTTTAAGAATTTTACTGTCTTC

#6: CTCGAGGGATCTAATCTTGATCATTTCTTCTTC

#7: CACAGTGGATGTTCCTGAGGATTTGAG

#8: CATTTCTGGACATAAGCATCAACTTCGTAC

CHAPTER 4*

FRG1 IS AN ACTIN-BUNDLING PROTEIN

* Some of the data and writing is adapted from

Chia-Yun Jessica Sun, Silvana van Koningsbruggen, Kirsten Straasheijm, Rinse Klooster, Steven W. Long, Takako I. Jones, Michel Bellini, Lyne Levesque, William M. Brieher, Silvère M. van der Maarel, and Peter L. Jones. ***FSHD region gene 1 (FRG1) is a dynamic RNA-associated, actin bundling protein.*** Submitted for publication.

INTRODUCTION

FRG1 is one of the leading candidate genes for FSHD. However, its relevance to FSHD pathology is unclear due to the lack of understanding of its molecular functions and inconsistent reports of mis-regulations in FSHD patient samples compared to controls (Arashiro et al., 2009; Gabellini et al., 2002; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; Winokur et al., 2003). In previous chapters, we have shown that FRG1 is a dynamic nuclear, cytoplasmic, and sarcomeric protein with distinct subcellular localizations. In regards to RNA-biogenesis, we found that endogenous FRG1 localized to the nucleoli and Cajal bodies, sites where RNA-biogenesis takes place. Further, FRG1 associates with RNA *in vivo* and binds directly to mRNA transcripts *in vitro*. Using immunostaining, we also identified for the first time, the cytoplasmic pool of human and mouse endogenous FRG1 in multiple cell types. This chapter will focus on the cytoplasmic functions of FRG1.

In silico analysis of FRG1 revealed two N-terminal NLSs, a carboxyl bipartite NLS sequence, and a central fascin-like domain, found in the fascin family of actin-bundling and cross-linking proteins (Kureishy et al., 2002). Our lab has recently characterized the *C. elegans* FRG1 homolog, FRG-1, and, in addition to the conserved strong nucleolar localization, we found a cytoplasmic pool of FRG-1 localized to body-

wall muscle dense bodies, structures functionally analogous to vertebrate Z-discs and costameres (Liu et al., 2010). This striking Z-disc localization for FRG1 is conserved in mouse and human skeletal muscle (M. L. Hanel, submitted for publication). Functionally, we found FRG-1 is an actin-binding and bundling protein *in vitro* suggesting FRG-1 may be a part of or involved in stabilization of the actin cytoskeleton (Liu et al., 2010). The actin-binding and bundling activity is also conserved in human FRG1 (Liu et al., 2010). These data suggest a role for vertebrate FRG1 in muscle development and function, possibly through the regulation of the actin cytoskeleton, and make FRG1 an intriguing candidate gene for FSHD.

In this chapter, we further investigated FRG1's function in the cytoplasm and its association with the cytoskeleton. We characterized human FRG1 actin binding activity and showed that FRG1 cytoplasmic localization is dependent upon the integrity of the actin cytoskeleton.

RESULTS

Human FRG1 binds F-actin and forms multimers in vitro

Recently, we have shown that the *C. elegans* and human FRG1 bind to and bundle F-actin (Liu et al., 2010). In order to further analyze the actin binding activity of human FRG1, a high-speed cosedimentation assay using recombinant human FRG1 was performed (Figure 4.1). Increasing concentrations of FRG1 were incubated with a consistent amount of F-actin and subjected to high-speed centrifugation to pellet the actin-associated FRG1. The ratios of FRG1 in the pellets (bound to actin) and supernatants (not bound to actin) were determined by SDS-PAGE and Coomassie BBR staining followed

by analysis with QuantityOne software. The results showed that human FRG1 binds actin with a Kd = 0.16 ± 0.03 μM and Bmax = 1.54 ± 0.02 (corresponding to two FRG1 per actin monomer) (Figure 4.1A). To determine if FRG1 could form the multimers required for F-actin cross-linking, a glutaraldehyde cross-linking assay using recombinant FRG1 was performed. A cross-linking time-course assay showed that FRG1 forms dimers and tetramers *in vitro* (Figure 4.1B), providing the multiple actin binding sites required for F-actin cross-linking.

Endogenous FRG1 interacts with actin cytoskeleton in C2C12 cells

Since FRG1 is an actin binding protein, potentially FRG1 associates with the actin cytoskeleton *in vivo*. Murine C2C12 cells subjected to immunostaining for FRG1 shows granular cytoplasmic staining (Figure 4.2A-D, Q-S). To determine if cytoplasmic FRG1 potentially associates with the actin cytoskeleton, as visualized by phalloidin staining, we treated the cells with various concentration of Latrunculin B (LatB) for 5 or 15 minutes. LatB inhibits actin polymerization and therefore disrupts actin cytoskeleton as indicated by loss of phalloidin staining (Figure 4.2A, E, I, M). The intensity of the cytoplamic pool of FRG1 decreased gradually as the intensity of LatB treatment increased, and FRG1 staining was abolished in phalloidin weak areas (Figure 4.2E-P) indicating that FRG1's localization in the cytoplasm is dependent on the intact actin cytoskeleton *in vivo*.

DISCUSSION

The nuclear localization of FRG1 seems to be essential considering its double NLS and bipartite sequences; yet despite these signals, a consistent cytoplasmic pool of FRG1

remains. This suggests that FRG1 is actively being retained in the cytoplasm and that the nuclear function of FRG1 is only part of the story. We have previously proposed that FRG1 has a role beyond the nucleus as well. The *C. elegans* FRG-1 bundles F-actin *in vitro* and localizes to the body wall muscle dense body, a structure analogous to the vertebrate muscle Z-disc and costameres (Liu et al., 2010). Here we show that human FRG1 retains the conserved actin binding activity, binding to actin in a ratio of 2:1, similar to its *C. elegans* homolog, and forms dimers capable of bundling F-actin (Liu et al., 2010). Thus, FRG1 may have a structural role in stabilizing the actin cytoskeleton or may be mediating other cellular functions by associating with the actin cytoskeleton. In support of the latter situation, we showed that FRG1 requires the intact actin cytoskeleton for maintaining its particular punctate cytoplasmic localization suggesting that actin filaments actively anchor FRG1 in the cytoplasm. Together with our findings in Chapter 3 that FRG1 is involved in RNA-biogenesis and may be mediating mRNA export through the TAP/NXT1 pathway; we propose that the nuclear and cytoplasmic functions are linked through RNA. FRG1 may be involved in mRNA export and anchor mRNA at its designated location through associating with actin cytoskeleton for localized translation. This could also be related to the skeletal muscle Z-disc localization of FRG1 in mouse and human, as FRG1 anchors the mRNA transcripts at the Z-disc for localized translation. Interestingly, several Z-disc proteins have been shown to co-localize with their cognate mRNAs in cultured skeletal muscle (Morris and Fulton, 1994). Another important aspect of Z-disc localization of FRG1 is that disruptions of numerous Z-disc proteins are causally related to myriad myopathies (McNally and Pytel, 2007; Selcen and Carpen, 2008). Z-disc is the essential sarcomeric structure to ensure proper aliments of the sarcomeres,

therefore maintains muscle integrity during contraction. The actin binding and bundling activities of FRG1 could also be important as it could be part of the structural proteins at the Z-disc.

Our data suggest FRG1 is a multifunctional protein with distinct subcellular localizations and the subcellular distribution of FRG1 may be highly regulated. Even small alterations of FRG1 protein levels could change the subcellular distribution of FRG1, potentially dysregulating its function, and ultimately leads to disruption of muscle development or integrity.

MATERIALS AND METHODS

Cell culture

C2C12 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 2 mM L-glutamine, supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin (pen/strep). Cells were incubated with 5% CO_2 at 37°C. For Latrunculin B (LatB) treatments, C2C12 cells were treated with indicated concentrations of LatB (Invitrogen Corp., Carlsbad, CA) for 5 or 15 min as indicated. After Lat B treatment, cells were fixed with 2% FA for 15 min at RT and subjected to immunostaining described below.

Immunocytochemistry

C2C12 cells were fixed in 2% formaldehyde (FA) in PBS for 15 min at room temperature (RT). After fixation, cells were made permeable with 0.25% Triton-X 100 in PBS for 10 min on ice, and subsequently blocked with 2% BSA in PBS for 30 min at RT.

Primary antibody incubation was carried out overnight at 4°C in PBS with 2% BSA, and secondary antibody incubation was for 40 min at RT in PBS. Secondary antibodies used were pre-cleared FITC-conjugated donkey anti-rabbit (Jackson ImmunoResearch Laboratories Inc) used at 1:100. F-actin was visualized with 5 U/ml rhodamine-phalloidin (Invitrogen Corp) incubated for 30 min at RT. The 4,6-diamidino-2-phenylindole (DAPI) was used at 1μg/ml.

Actin binding assay

Actin was purified from rabbit skeletal muscle as described (Pardee and Spudich, 1982). Recombinant full-length FRG1 generated as described in Chapter 3 was further purified by ion exchange chromatography, binding to MonoS resin using an AKTA-FPLC (GE Healthcare), step eluted between 350 and 700 mM NaCl in 10 mM HEPES pH 7.5 and 10% glycerol, and dialyzed to 50 mM NaCl. Actin co-sedimentation assays were carried out as described with minor modifications (Ono et al., 1997). For high-speed sedimentation assays, increasing amounts of recombinant FRG1 (7-30 mM as monomer) were incubated with G-actin (final concentration 6 μM) in 50 mM NaCl in a total volume of 50 μl at RT for 2 hrs and then centrifuged at 100,000 *g* for 20 min at RT. The amounts of FRG1 and F-actin in the supernatants and pellets were determined by densitometry using Coomassie-Blue-stained 10% SDS-PAGE. The intensities of the stained polypeptide bands were quantified by volume integration after local background subtraction using Bio-Rad Quantity One software (Bio-Rad Laboratories, Hercules, CA). Subtracting the percentage of FRG1 in the pellet without actin was used to normalize the FRG1. The binding data from four independent experiments were analyzed by fitting to

the Michaelis-Menton equation, Y=Bmax*X/(Kd+X), using the nonlinear regression function of Prism 5 (GraphPad Software), where Y is FRG1 per F-actin (mol/mol) in the pellet and X is the FRG1 concentration (micromolar) remaining in the supernatant.

Glutaraldehyde cross-linking

Recombinant full-length FRG1 protein (0.1 mg/ml) was incubated with 0.01% glutaraldehyde in 1X PBS in 1 ml at RT. Samples (50 µl) were removed at the indicated time intervals and the reactions were stopped by the addition of 2X Laemmli SDS-PAGE buffer, boiled, and separated by SDS-PAGE (12%). Protein was detected by western blotting using the HS2 antibody described above.

Microscopy

Standard fluorescence microscopy was carried out using a HCL FL Fluotar 100X oil objective (NA=1.30) on an upright Leica DMR microscope. Images were captured using a monochrome Retiga EXI Charge-Coupled Device (CCD) camera (Qimaging) driven by the In vivo software (version 3.2.0, Media Cybernetics). All images were processed using Adobe Photoshop to adjust brightness, contrast, size, and merged or split channels. Applied Precision Personal Deltavision was used for deconvolution images.

REFERENCES

Arashiro, P., I. Eisenberg, et al. 2009. Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. Proc Natl Acad Sci U S A, **106**: 6220-5.

Gabellini, D., M. R. Green, et al. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell, **110**: 339-48.

Jiang, G., F. Yang, et al. 2003. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet, **12**: 2909-21.

Klooster, R., K. Straasheijm, et al. 2009. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. Eur J Hum Genet, **17**: 1615-24.

Kureishy, N., V. Sapountzi, et al. 2002. Fascins, and their roles in cell structure and function. Bioessays, **24**: 350-61.

Liu, Q., T. I. Jones, et al. 2010. Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. J Cell Sci, **123**: 1116-23.

Masny, P. S., O. Y. Chan, et al. 2010. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. Eur J Hum Genet, **18**: 448-56.

McNally, E. M. and P. Pytel. 2007. Muscle diseases: the muscular dystrophies. Annu Rev Pathol, **2**: 87-109.

Morris, E. J. and A. B. Fulton. 1994. Rearrangement of mRNAs for costamere proteins during costamere development in cultured skeletal muscle from chicken. J Cell Sci, **107 (Pt 3)**: 377-86.

Ono, S., Y. Yamakita, et al. 1997. Identification of an actin binding region and a protein kinase C phosphorylation site on human fascin. J Biol Chem, **272**: 2527-33.

Osborne, R. J., S. Welle, et al. 2007. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology, **68**: 569-77.

Pardee, J. D. and J. A. Spudich. 1982. Purification of muscle actin. Methods Enzymol, **85 Pt B**: 164-81.

Selcen, D. and O. Carpen. 2008. The Z-disk diseases. Adv Exp Med Biol, **642**: 116-30.

Winokur, S. T., K. Barrett, et al. 2003. Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. Neuromuscul Disord, **13**: 322-33.

FIGURES

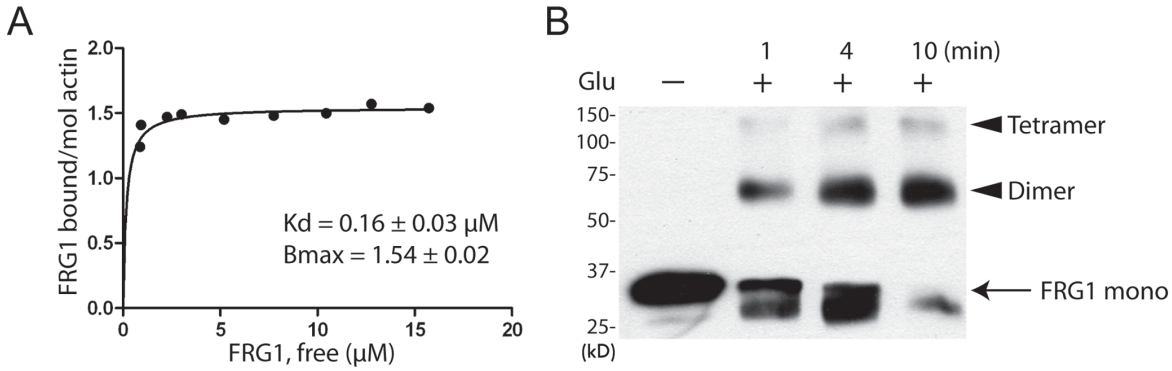

Figure 4.1: Human FRG1 binds F-actin and forms multimers *in vitro*. (A) High speed cosedimentation assay using an increasing amount of recombinant human FRG1 with a constant concentration of F-actin shows saturation. (B) Western blot analysis of a glutaraldehyde crosslinking time-course assay using recombinant human FRG1 shows FRG1 can readily form dimers and tetramers *in vitro*.

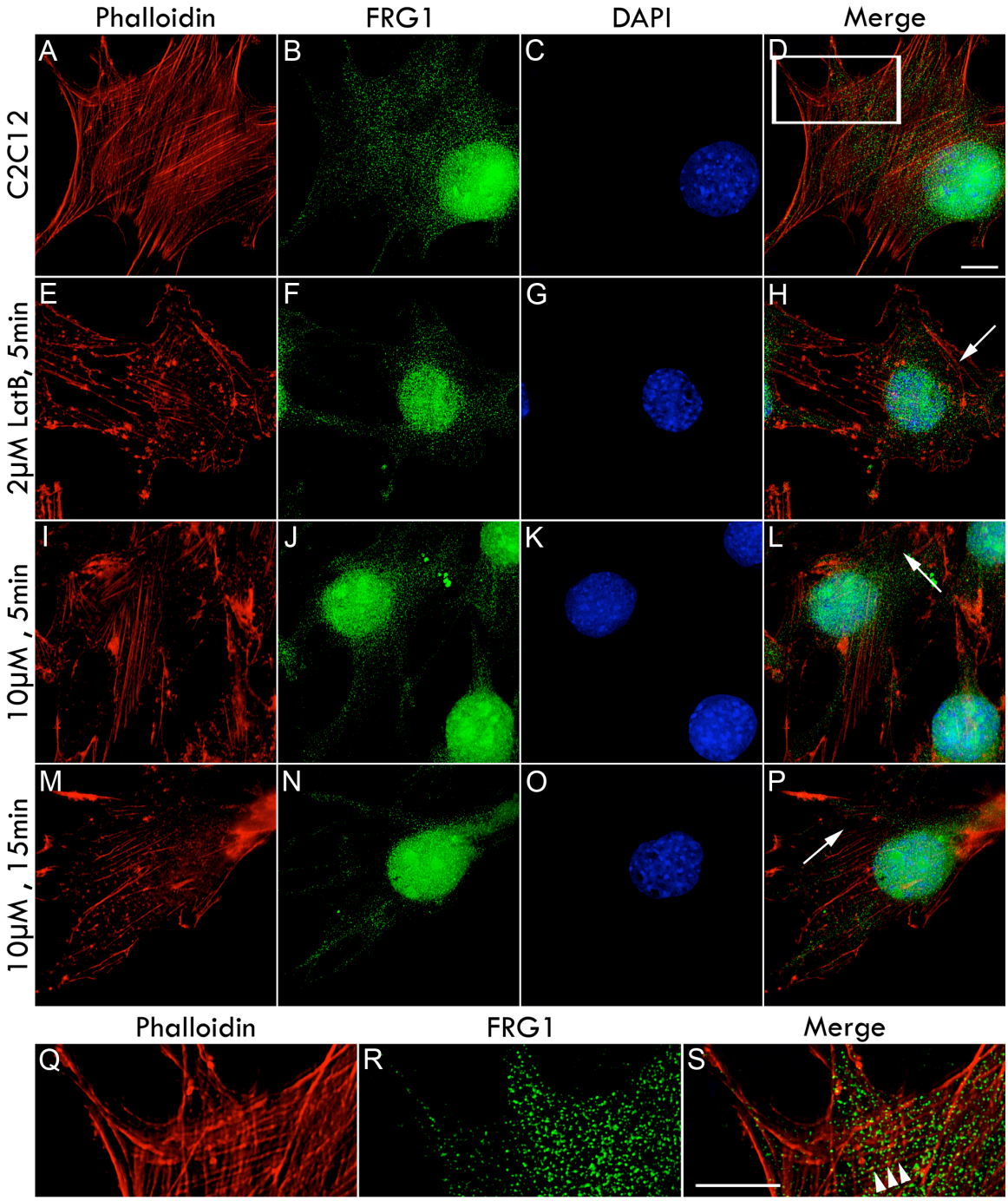

Figure 4.2: Endogenous FRG1 requires the actin cytoskeleton for cytoplasmic localization in murine C2C12 cells. (A-D, Q-S) Cytoplasmic pool of FRG1 appeared granular and associated with actin filaments labeled by rhodamine-phalloidin staining (arrowheads in S). Boxed area in D is shown magnified 2X in Q-S. (E-P) Cells were treated with variable concentrations of Latrunculin B (LatB) for 5 or 15 minutes as indicated. FRG1 staining decreased or was abolished (arrows) in areas of LatB disrupted actin filaments as indicated by weak phalloidin staining.

CHAPTER 5

TRANSLATIONAL IMPACT TOWARD FSHD

This thesis aims to understand the normal cellular function of the human FRG1 protein, considered one of the leading candidates for mediating FSHD pathophysiology. Even though the genetic lesion of FSHD was mapped twenty years ago to the D4Z4 repeat array on chromosome 4q35, the subtelemeric region of chromosome 4q (Wijmenga et al., 1990; Wijmenga et al., 1992), the molecular mechanism of the disease remains elusive as how the contraction of the array leads to FSHD pathology. It is unclear whether the repeat element itself contains the gene responsible for disease pathogenesis (*DUX4*), or the contraction causes a downstream effect on other candidate genes in a *trans* or *cis* fashion, or the disease is a combination of both mechanisms. Expression profile studies and analyses failed to show consistent mis-regulation of any candidate gene in patient samples compared to normal individuals (Dixit et al., 2007; Gabellini et al., 2002; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; Winokur et al., 2003). This can partially be explained by the differences in the samples collected, processing procedures, or the techniques used for analysis. Muscle biopsy that appropriately represents the disease progression is hard to collect as the affected muscles often show fibrotic or adipose infiltration, and biopsies collected are usually mixed with multiple cell types such as epithelial cells. Thus, in order to understand the molecular mechanism of FSHD pathogenesis, it is important to investigate the normal cellular functions of the candidate genes and see how its mis-regulation would contribute to FSHD pathology. Our work here provides insights into the subcellular localizations and

biological activities of FRG1, and shows how mis-regulation of FRG1 can potentially be detrimental to muscle development or integrity. Further, these activities would have additional interesting implications beyond their significance in muscle development should FRG1 be determined to be involved in mediating FSHD pathophysiology.

We have previously shown that FRG1 is critical for vertebrate musculature and vasculature development; here we have addressed the potential for FRG1 misexpression to lead to a muscular dystrophy phenotype by understanding its normal cellular function. We have further investigated the subcellular localizations of endogenous FRG1 and discovered both nuclear and cytoplasmic pools of FRG1 in multiple cell types. The distribution of FRG1 is consistent between most cell types with very prominent nuclear staining (granular staining in the nucleoplasm and enriched in nucleoli) followed by granular staining in the cytoplasm, except for undifferentiated and fully differentiated HSMM (human skeletal muscle myoblasts) showing the greatest amount in the cytoplasm. The subcellular localization greatly changes during myogenesis with increased nuclear and nucleolar staining in early myogenesis, and eventually localized to the Z-disc in mature myotubes. This is consistent with previous finding that FRG1 level is upregulated during myogenic differentiation in human myoblasts (Bodega et al., 2007) (T.I. Jones, unpublished observation), and the overexpressed FRG1 is predominantly nuclear and nucleolar (van Koningsbruggen et al., 2004). The Z-disc localization in mature HSMM myotubes is in agreement with our finding in *C. elegans* that FRG-1 (FRG1 ortholog) localizes to the body wall dense bodies, which is functionally analogous to the vertebrate Z-discs and costameres (Moerman and Williams, 2006). This is the first link between a FSHD candidate gene and the sarcomeric phenotype of FSHD, where the structures at the

sarcolemma are misaligned and the attachment between the sarcolemma and the underlying contractile are disrupted in FSHD patient samples (Reed et al., 2006). Therefore, FRG1 is similarly localized with many other proteins associated with myopathies, strongly implicating its involvement in FSHD.

Focusing on the nuclear function of FRG1, here we show endogenous FRG1 localizes to Cajal bodies (CBs) and the granular component (GC) of nucleoli, as well as associated with nascent mRNA transcripts and the actively transcribed regions of the euchromatin. CBs are nuclear bodies that store factors involved in transcription and RNA processing, and sites for snRNP biogenesis (Austin and Bellini, 2010; Gall, 2003). Nucleoli are sites for transcription assembly of pre-ribosomal particles, maturations of some RNP particles, and assembly of the RNA splicing machinery (Austin and Bellini, 2010; Boisvert et al., 2007). Lampbrush chromosome spreads from *X. laevis* oocytes confirmed localization of endogenous FRG1 in CBs and nucleoli, and further indicate FRG1 associate with nascent RNA polymerase II generated mRNAs *in vivo*. RNA-IP from HeLa cells also showed FRG1 associates with mRNA transcripts *in vivo*, and the association may be direct as FRG1 bind mRNA transcripts *in vitro* through its N-terminal NLS. Together, these data further indicates FRG1 is involved in RNA-biogenesis or pre-mRNA processing. Thus, FRG1 is a nuclear and cytoplasmic protein associated with the contractile machinery and RNA transcripts.

Our collaborative work with Silvére van der Maarel's group in Leiden, the Netherlands, has led to the identification of two FRG1 protein binding partners, karyopherinα2 and TAP. Karyopherinα2, a nuclear import protein, is a prominent FRG1 binding partner by yeast-two-hybrid screens. A direct interaction was confirmed by GST

pull-down and co-IP experiments (Sun et al., submitted for publication). Although proteins such as FRG1 (30 kb) are small enough to pass through the nuclear pore by diffusion, active nuclear transport via the nuclear import receptor karyopherinα2 results in rapid and efficient nuclear accumulation, which likely indicates an essential nuclear function for FRG1. The N-terminal NLS region (contains two NLS sequences, NLS1 and NLS2), not the carboxyl bipartite NLS, was demonstrated to be the functional NLS for FRG1 (van Koningsbruggen et al., 2004). We further mapped NLS2 as the essential NLS and showed that the subcellular localization of FRG1 can potentially be regulated through phosphorylation to ensure a spatial-temporal specific distribution of FRG1 (Sun et al., submitted for publication). TAP, an mRNA export receptor, is a new binding partner of FRG1, suggesting FRG1 may involve in mediating mRNA export through TAP/NXT1 pathway. Interestingly, FRG1 interacts directly with TAP through the N-terminal NLS, the functional NLS for FRG1 and the potential RNA binding domain. FRG1 may interact with TAP/NXT1 heterodimer and mediate mRNA export, and continue to associate with mRNA transcripts in the cytoplasm, therefore masking the NLS and maintain its cytoplasmic localization. Thus, FRG1 is a sarcomeric protein potentially involved in mRNA trafficking.

Further, we investigated the cytoplasmic function of and showed that human FRG1 bundles actin *in vitro* and may potentially be regulated through phosphorylation of two potential phosphorylation sites within the fascin like domain (Appendix A). Considering the associations of FRG1 with mRNA biogenesis and trafficking, we propose a model whereby FRG1's nuclear and cytoplasmic functions are linked through RNA (Figure 5.1). FRG1 may associate with mRNA transcripts in the nucleus, which are then

exported together to the cytoplasm, mediated through the TAP/NXT1 pathway, where the actin cytoskeleton anchors the FRG1-mRNA complex at its designated location. This is in agreement with our finding that the nuclear export of FRG1 is dependent on RNA, as inhibiting RNA polymerase II results in the accumulation of FRG1 in the recipient nuclei; and intact actin cytoskeleton is required for the granular staining of FRG1 in the cytoplasm of murine C2C12 cells.

Localized translation of mRNA has emerged as a new mechanism for regulating cellular protein synthesis in a spatial-temporal manner. In this model, mRNA transcripts are associated with repressor complexes and are silenced during transport until they reach their final destination. Translational repressors prevent protein synthesis by targeting different regulators of the translation process; the most frequently regulated step is translation initiation, the general rate-limiting step (Besse and Ephrussi, 2008). Translation of targeted mRNAs is activated upon the arrival at its subcellular destination, in response to spatial cues. Or, the targeted mRNAs are stored in a dormant state at their designated location, and is only activated in response to specific external signals (Besse and Ephrussi, 2008). Translational derepression might commonly be mediated by decreasing the affinity of translational repressors with the target mRNAs, and allow the initiation of translation to proceed.

FRG1 may potentially be involved in this process as FRG1 is deposited on mRNAs in the nucleus, mediating its export through the TAP/NXT1 pathway, anchoring the targeted mRNAs at its final locations in a repressed state, waiting for activation signals in response to external stress. We have found FRG1 localizes to Z-disc the in mature myotubes, and interestingly, several Z-disc proteins have been shown to co-localize with

their cognate mRNAs in cultured skeletal muscle (Morris and Fulton, 1994). Further, various translational regulators that are found in transport-RNP complexes are shuttling proteins that contain nuclear localization signals and accumulate at least transiently in the nucleus (Besse and Ephrussi, 2008). Finally, localizing mRNAs seem to co-sediment poorly with fractions containing actively translated mRNAs (polysomal fractions), which is consistent with our finding that FRG1 does not associate with actively translated polyribosomes (unpublished observation).

FRG1 is clearly an important multifunctional protein involved in muscle development, however a role in FSHD pathophysiology is controversial due to the failure to consistently find any significant changes in FRG1 levels between patients and unaffected individuals. Still, recent data towards understanding FRG1 is compelling. Overexpression of FRG1 specifically disrupts development of the vertebrate musculature and vasculature, the two tissues most affected in FSHD (Hanel et al., 2009; Wuebbles et al., 2009). Cytoplasmic FRG1 localizes to the skeletal muscle Z-disc in mouse and humans, the subcellular localization of numerous proteins related to myriad myopathies (McNally and Pytel, 2007; Selcen and Carpen, 2008). We subsequently added to the circumstantial evidence supporting FRG1's potential in FSHD by showing FRG1 is a dynamic RNA-associated actin binding protein in consistent with previous finding that RNA biogenesis is disrupted in FSHD (Davidovic et al., 2008; Gabellini et al., 2006). In addition, FRG1's interactions with both a nuclear importer and a nuclear exporter indicate its subcellular localization is highly regulated. Interestingly, in all systems tested including mammalian cell culture, *C. elegans*, *Xenopus*, and *Drosophila*, overexpressed FRG1 preferentially accumulates in the nucleus (Liu et al., 2010; van Koningsbruggen et

al., 2004; van Koningsbruggen et al., 2007) (PL Jones unpublished observation). Even small alterations of FRG1 protein levels could change the subcellular distribution of FRG1, potentially dysregulating its function. Increasing nuclear levels of FRG1 may ultimately result in mis-spliced mRNA transcripts, altered mRNA stability, or affect mRNA transport, any of which could adversely affect maintenance of muscle integrity.

REFERENCES

Austin, C. M. and M. Bellini. 2010. The dynamic landscape of the cell nucleus. Mol Reprod Dev, **77**: 19-28.

Besse, F. and A. Ephrussi. 2008. Translational control of localized mRNAs: restricting protein synthesis in space and time. Nat Rev Mol Cell Biol, **9**: 971-80.

Bodega, B., M. F. Cardone, et al. 2007. Evolutionary genomic remodelling of the human 4q subtelomere (4q35.2). BMC Evol Biol, **7**: 39.

Boisvert, F. M., S. van Koningsbruggen, et al. 2007. The multifunctional nucleolus. Nat Rev Mol Cell Biol, **8**: 574-85.

Davidovic, L., S. Sacconi, et al. 2008. Alteration of expression of muscle specific isoforms of the fragile X related protein 1 (FXR1P) in facioscapulohumeral muscular dystrophy patients. J Med Genet, **45**: 679-85.

Dixit, M., E. Ansseau, et al. 2007. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. Proc Natl Acad Sci U S A, **104**: 18157-62.

Gabellini, D., G. D'Antona, et al. 2006. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. Nature, **439**: 973-7.

Gabellini, D., M. R. Green, et al. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell, **110**: 339-48.

Gall, J. G. 2003. [A role for Cajal bodies in assembly of the nuclear transcription machinery]. Tsitologiia, **45**: 971-5.

Hanel, M. L., R. D. Wuebbles, et al. 2009. Muscular dystrophy candidate gene FRG1 is critical for muscle development. Dev Dyn, **238**: 1502-12.

Jiang, G., F. Yang, et al. 2003. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet, **12**: 2909-21.

Klooster, R., K. Straasheijm, et al. 2009. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. Eur J Hum Genet, **17**: 1615-24.

Liu, Q., T. I. Jones, et al. 2010. Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. J Cell Sci, **123**: 1116-23.

Masny, P. S., O. Y. Chan, et al. 2010. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. Eur J Hum Genet, **18**: 448-56.

McNally, E. M. and P. Pytel. 2007. Muscle diseases: the muscular dystrophies. Annu Rev Pathol, **2**: 87-109.

Moerman, D. G. and B. D. Williams. 2006. Sarcomere assembly in C. elegans muscle. WormBook: 1-16.

Morris, E. J. and A. B. Fulton. 1994. Rearrangement of mRNAs for costamere proteins during costamere development in cultured skeletal muscle from chicken. J Cell Sci, **107 (Pt 3)**: 377-86.

Osborne, R. J., S. Welle, et al. 2007. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology, **68**: 569-77.

Reed, P., N. C. Porter, et al. 2006. Sarcolemmal reorganization in facioscapulohumeral muscular dystrophy. Ann Neurol, **59**: 289-97.

Selcen, D. and O. Carpen. 2008. The Z-disk diseases. Adv Exp Med Biol, **642**: 116-30.

van Koningsbruggen, S., R. W. Dirks, et al. 2004. FRG1P is localised in the nucleolus, Cajal bodies, and speckles. J Med Genet, **41**: e46.

van Koningsbruggen, S., K. R. Straasheijm, et al. 2007. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. Chromosoma, **116**: 53-64.

Wijmenga, C., R. R. Frants, et al. 1990. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. Lancet, **336**: 651-3.

Wijmenga, C., J. E. Hewitt, et al. 1992. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. Nat Genet, **2**: 26-30.

Winokur, S. T., Y. W. Chen, et al. 2003. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. Hum Mol Genet, **12**: 2895-907.

Wuebbles, R. D., M. L. Hanel, et al. 2009. FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. Dis Model Mech, **2**: 267-74.

FIGURES

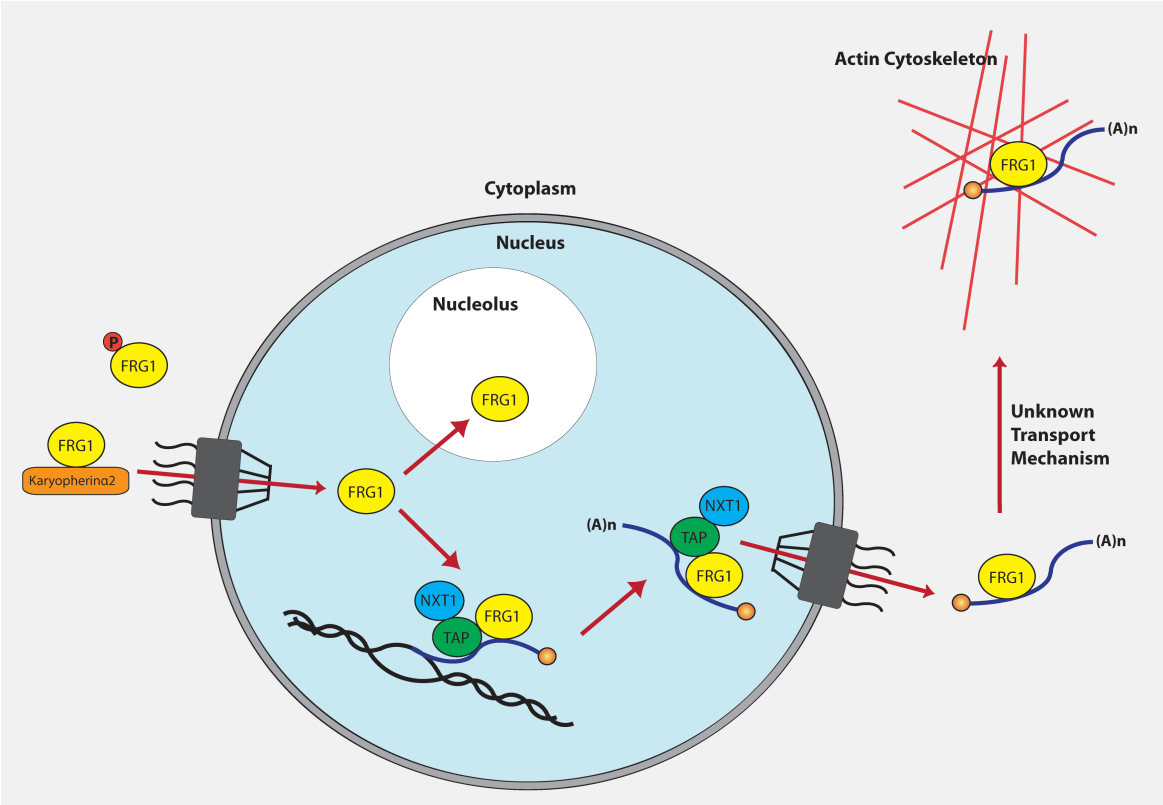

Figure 5.1: Proposed model illustrating FRG1 activities. FRG1, translated in the cytoplasm, is transported to the nucleus via the nuclear importer receptor karyopherinα2 in a phosphorylation-dependent manner. Once in the nucleus, FRG1 either translocates to the nucleolus or associates with nascent mRNA transcripts that are exported together through the TAP/NXT1 pathway. In the cytoplasm, the FRG1-mRNA complexes are transported through an unknown mechanism and anchored at its final location by the actin cytoskeleton.

APPENDIX A

PHOSPHORYLATION OF HUMAN FRG1 AFFECTS ITS ACTIN BUNDLING ACTIVITY

INTRODUCTION

FRG1 in one of the leading candidate genes for FSHD, however, its involvement in FSHD pathophysiology is unclear and its precise function unknown. FRG1 has multiple NLS sequences and has a proposed role in RNA-biogenesis (van Koningsbruggen et al., 2004; van Koningsbruggen et al., 2007). Further, we have shown that FRG1 potentially is involved in mRNA transport and regulation. FRG1 contains a central fascin-like domain, which is found in actin bundling or cross-linking proteins (Kureishy et al., 2002) and we have shown that human FRG1 can bind actin (two FRG1s per actin monomer) and bundle F-actin *in vitro* and requires F-actin for its cytoplasmic localization *in vivo*. Since FRG1 only has one fascin-like domain, and therefore likely only one actin-binding domain, FRG1 would require multimerization for cross-linking actin. In fact, FRG1 can associate with actin through multimerization, as FRG1 forms homodimers and homotetramers *in vitro* as shown by a glutaraldehyde cross-linking assay. Here, we further characterize FRG1's actin binding and bundling.

The association of FRG1 with the actin cytoskeleton was discussed in Chapter 4; here we want to investigate how this association can be regulated. We made a series of deletions and several point mutations within the fascin-like domain of FRG1, generated recombinant protein, and mapped the regions important for actin bundling activity using an *in vitro* fluorescent F-actin bundling assay. Three separate regions within the FRG1 fascin-like domain appeared to mediate the actin bundling activity. Interestingly, two

potential highly conserved protein kinase C (PKC) phosphorylation sites were found in two of these regions and several actin bundling proteins, including fascin, are regulated by PKC phosphorylation in their actin binding sites. We further showed that FRG1 can be phosphorylated *in vitro*, and mutation of the phosphorylation sites abolishes the actin bundling activity of FRG1. These data suggest FRG1's actin bundling activity is regulated by phosphorylation.

RESULTS

Three regions within FRG1's fascin-like domain are required for F-actin bundling activity

A series of internal FRG1 deletion constructs, each containing a deletion of 15-20 amino acids, spanning the fascin-like domain, were made in order to map the region of FRG1 that mediates its F-actin bundling activity (Figure A.1). A fluorescent F-actin bundling assay was carried out by incubating the recombinant FRG1 deletion protein of interest with TMR-labeled F-actin, which allows the bundled F-actin be readily visualized under a fluorescent microscope (Figure A.2). Using this assay, three separate regions, Δ72-86, Δ103–118, and Δ134-177, were found to be required for the actin bundling activity of FRG1 (Figure A.3).

FRG1 can be phosphorylated in vitro by PKC

Interestingly, *in silico* sequence analysis of the predicted FRG1 amino acid sequence identified two potential PKC phosphorylation sites, FTL and KSG, within the fascin-like domain. These sites are 100% evolutionarily conserved from human through

Drosophila and each resides in a distinct region that is required for the actin bundling activity (Figure A.4). To determine whether these PKC sites are potential substrates for PKC we performed an *in vitro* phosphorylation assay using recombinant PKC and [^{32}P] γ-ATP on recombinant full-length FRG1 or FRG1 point mutation constructs where the FTL and/or KSG residues were changed to alanines (A). We found that the full-length FRG1 was phosphorylated by PKC and the phosphorylation level decreased significantly upon mutation of each of the potential PKC sites (Figure A.5). Next, we want to see if the point mutations of the potential PKC sites would affect the actin bundling activity of FRG1. We performed the fluorescent actin bundling assay on the recombinant FRG1 with point mutations and showed mutation of the potential PKC sites abolishes the actin bundling activity of FRG1. Thus, FRG1's actin bundling activity potentially could be regulated by PKC.

DISCUSSION

FRG1 is a leading candidate gene for FSHD; however, its involvement in FSHD pathophysiology is not yet established due to its unknown function and inconsistent reports of its misregulation in FSHD samples compared to controls (Gabellini et al., 2002; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; Winokur et al., 2003). The field has focused on its nuclear function based on the initial subcellular localization studies of exogenously expressed FRG1 showing FRG1 is exclusively nuclear in cells overexpressing epitope-tagged FRG1 (van Koningsbruggen et al., 2004). We have shown that the endogenous FRG1 is abundant in the cytoplasm of all cells tested and the intact actin cytoskeleton is required for the granular cytoplasmic staining of FRG1 in

C2C12 cells (Chapter 4). Interestingly, we have also shown that FRG1 has an actin binding and bundling activity that is conversed in its *C. elegans* homolog, FRG-1 (Liu et al., 2010), consistent with a cytoplasmic role. This study focused on this actin bundling activity, mediated by FRG1's fascin-like domain. Fascins are a conserved group of actin-bundling proteins that stabilize actin filaments for cellular processes. Each vertebrate fascin contains four fascin-like domains, and two of which contain single actin-binding sites (Edwards and Bryan, 1995; Kureishy et al., 2002). FRG1 contains a single fascin-like domain, and we have previously shown that FRG1 forms homodimers *in vitro*, which provides the multiple actin-binding sites necessary for F-actin bundling. Here we have mapped three regions within the fascin-like domain that mediates the actin bundling activity of FRG1. These regions may be the dimerization or actin binding domains of FRG1, or regulate actin bundling activity through other mechanisms.

Two highly conserved potential PKC sites reside in two of the three regions mediating FRG1 actin bundling activity. We have shown both PKC sites are functional *in vitro*, and mutation of each PKC sites is sufficient to abolish the actin bundling activity of FRG1. Interestingly, the actin binding and bundling activities of human Fascin are also regulated through phosphorylation (Ono et al., 1997; Yamakita et al., 1996). Further experiments are required to investigate if FRG1 is phosphorylated *in vivo*, and how phosphorylation of FRG1 affects the actin binding activity.

MATERIALS AND METHODS

DNA constructs

All PCR primers are listed in Table A1. To generate the FRG1 His-tagged bacterial expression constructs, the cDNA for the full-length human FRG1 coding sequence was PCR amplified (primers #1 and #2) and subcloned between the NdeI and XhoI restriction sites into pET-23b vector (Novagen, Gibbstown, NJ). For the fascin deletion constructs, use Δ58-71 as example, two PCR reactions were generated using #1 and #3, #2 and #4, respectively. The PCR products were then used as templates for another PCR reaction using primer #1 and #2. The final PCR product was subcloned between the NdeI and XhoI restriction sites into pET-23b vector (Novagen, Gibbstown, NJ). For point mutations of potential phosphorylation sites, FTL and KSG, these sites were replaced by three alanines (A) using a similar technique as making the fasin domain deletions. For FTL site, two separate PCR were performed using primer sets #1 and #19, #2 and #20, respectively. The PCR products were subsequently used as templates for PCR reaction using primers #1 and #2, and the final product was subcloned between the NdeI and XhoI restriction sites into pET-23b vector. For KSG site, two separate PCR were performed using primer sets #1 and #21, #2 and #22, respectively. The PCR products were subsequently used as templates for PCR reaction using primers #1 and #2, and the final product was subcloned between the NdeI and XhoI restriction sites into pET-23b vector.

Fluorescent actin bundling assay

Actin was purified from rabbit skeletal muscle as described (Pardee and Spudich, 1982). To generate full-length and deletion-containing FRG1 recombinant proteins, the pET23 plasmid constructs described above were transformed into E. coli BL21 (DE3) cells and induced with 1 mM IPTG. The protein was purified using TALON resin (Clontech, Mountain View, CA) as per the manufacturer's instructions. Recombinant full-length FRG1 or deletion construct proteins were further purified by ion exchange chromatography, binding to MonoS resin using an AKTA-FPLC (GE Healthcare), eluted with gradient salt between 100 and 1000 mM NaCl and peak fractions containing FRG1 proteins are selected and used for the assay. A mixture of unlabeled and tetramethylrhodamine labeled G-actin (TMR-G-actin) in a ratio of 1:1 is prepared in G-buffer (2mM Tris pH 8.0, 0.2 mM ATP, 0.5 mM DTT, and 0.2 mM $CaCl_2$). Full-length and fascin domain deletions of FRG1 were incubated with G-actin and TMR-G-actin in bundling buffer (100 mM HEPES pH7.4, 50 mM KCl, 2 mM $MgCl_2$, 2 mM ATP, and 1 mM EGTA). The reaction was carried out in RT for 5 minutes. 3 µl of reaction was visualized under fluorescent microscope to detect F-actin bundles.

In vitro phosphorylation assay of FRG1

Full-length and point mutations of FRG1 are expressed in E. coli BL21 (DE3) and purified by ion exchange chromatography (binding to MonoS resin) as described above. The phosphorylation assay was carried out essentially as described with minor modifications (Yamakita et al., 1996). Full-length and point mutation constructs of FRG1 were incubated with Protein Kinase C (PKC, Enzo) in a final concentration of 1 µg/ml,

and [γ-^{32} P] labeled ATP (PerkinElmer) in phosphorylation buffer (30 mM KCl, 3 mM $MgCl_2$, 0.2 mM EGTA, and 20 mM Tris pH 7.5) in the presence of PMA (phorbal 12-myristate 13 acetate, a potent PKC activator, Enzo) and DTT in the final concentration of 100 ng/ml and 0.5 mM, respectively. The reaction was incubated for 1 hour at 30°C. 1.2 μg of full-length FRG1 or point mutation constructs were used in each reaction. After incubation, the reactions were analyzed by Coomassie-Blue-stained 12% SDS-PAGE, and the phosphorylation levels were visualized by autoradiography.

REFERENCES

Edwards, R. A. and J. Bryan. 1995. Fascins, a family of actin bundling proteins. Cell Motil Cytoskeleton, **32**: 1-9.

Gabellini, D., M. R. Green, et al. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell, **110**: 339-48.

Jiang, G., F. Yang, et al. 2003. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet, **12**: 2909-21.

Klooster, R., K. Straasheijm, et al. 2009. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. Eur J Hum Genet, **17**: 1615-24.

Kureishy, N., V. Sapountzi, et al. 2002. Fascins, and their roles in cell structure and function. Bioessays, **24**: 350-61.

Liu, Q., T. I. Jones, et al. 2010. Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. J Cell Sci, **123**: 1116-23.

Masny, P. S., O. Y. Chan, et al. 2010. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. Eur J Hum Genet, **18**: 448-56.

Ono, S., Y. Yamakita, et al. 1997. Identification of an actin binding region and a protein kinase C phosphorylation site on human fascin. J Biol Chem, **272**: 2527-33.

Osborne, R. J., S. Welle, et al. 2007. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology, **68**: 569-77.

Pardee, J. D. and J. A. Spudich. 1982. Purification of muscle actin. Methods Enzymol, **85 Pt B**: 164-81.

van Koningsbruggen, S., R. W. Dirks, et al. 2004. FRG1P is localised in the nucleolus, Cajal bodies, and speckles. J Med Genet, **41**: e46.

van Koningsbruggen, S., K. R. Straasheijm, et al. 2007. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. Chromosoma, **116**: 53-64.

Winokur, S. T., Y. W. Chen, et al. 2003. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. Hum Mol Genet, **12**: 2895-907.

Yamakita, Y., S. Ono, et al. 1996. Phosphorylation of human fascin inhibits its actin binding and bundling activities. J Biol Chem, **271**: 12632-8.

FIGURES AND TABLES

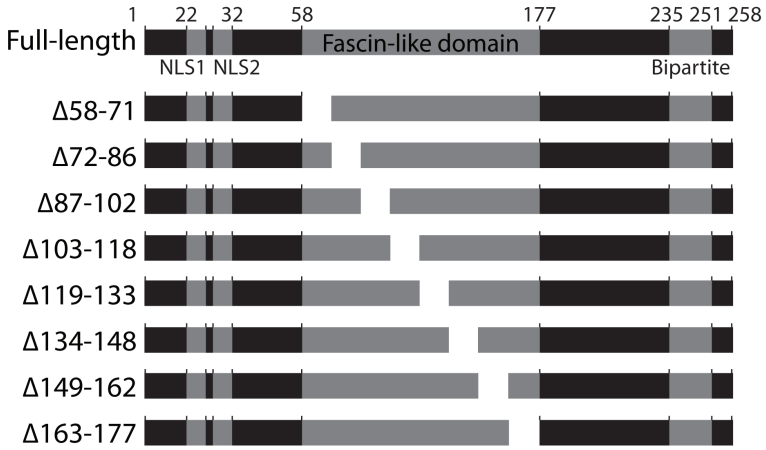

Figure A.1: Schematic view of FRG1 fascin domain deletions. A series of deletion constructs, each deleting 15-20 amino acids spanning the fascin-like domain, were made to map regions that mediates the actin bundling activity of FRG1.

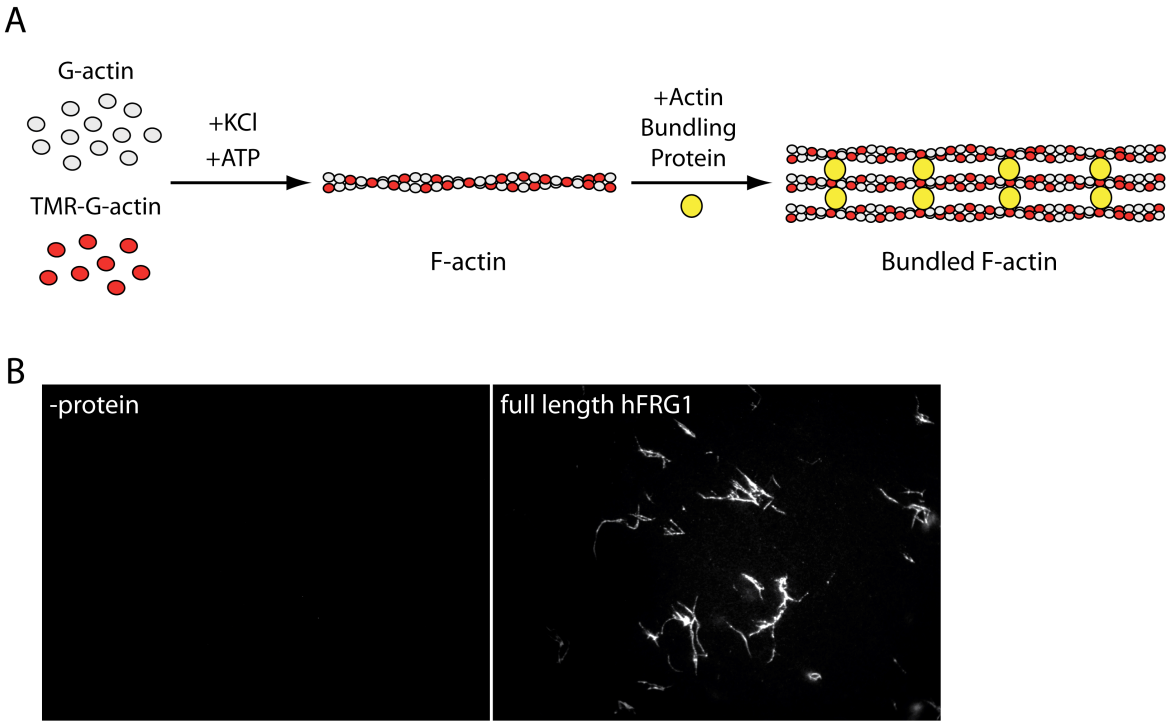

Figure A.2: Fluorescent actin bundling assay. Full-length FRG1 or fascin domain deletions were incubated with TMR-labeled F-actin, and visualized under fluorescent microscope for bundled F-actin.

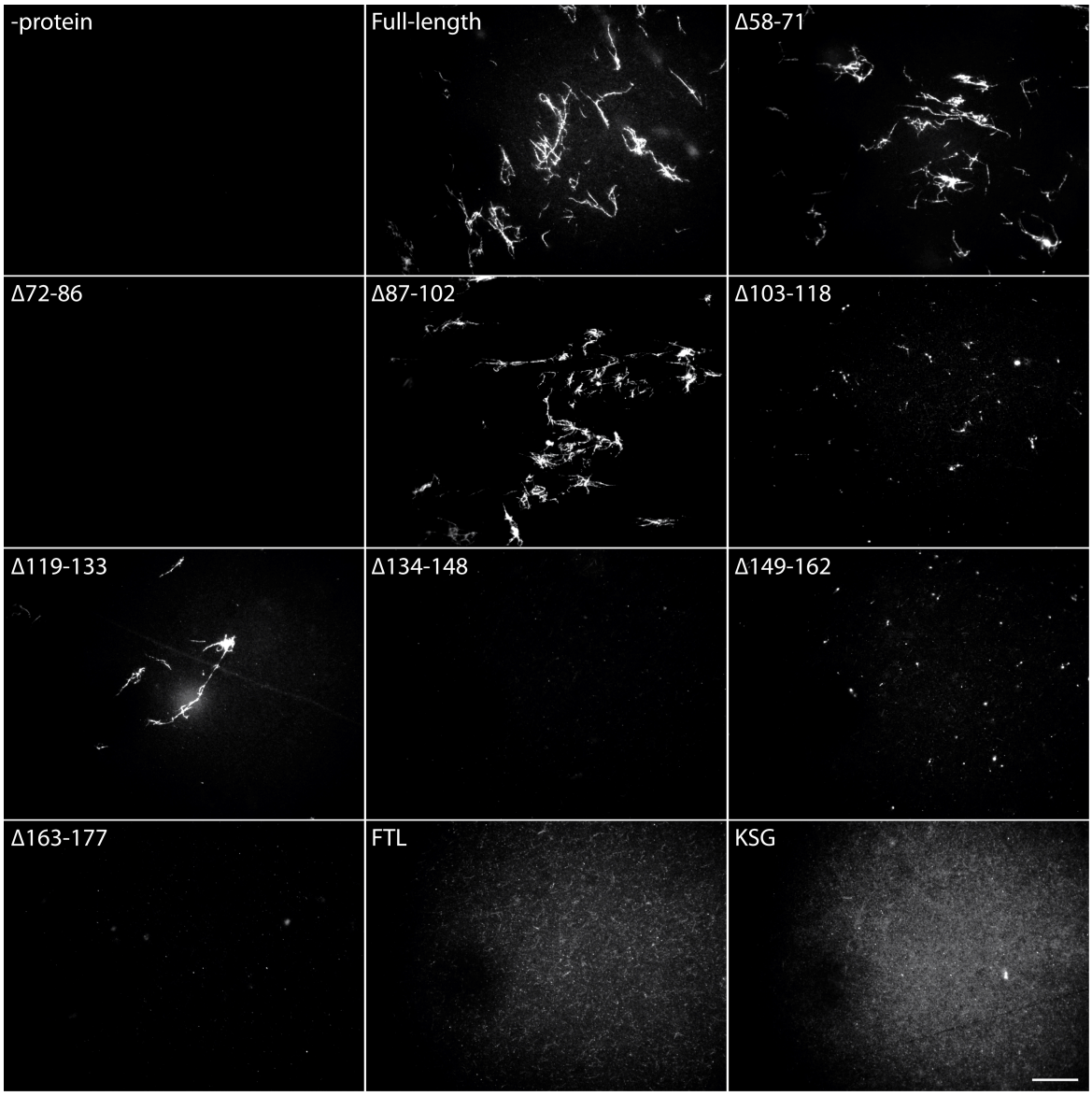

Figure A.3: Three separate regions within fascin-like domain are responsible for mediating actin bundling activity. Using deletion constructs in fluorescent bundling assay, we identified three regions responsible for actin bundling activity: Δ72-86, Δ103–118, and Δ134-177. Mutation of potential PKC sites abolishes the actin bundling activity of FRG1. Proteins used for fluorescent bundling assay are indicated and Bar = 10 μm.

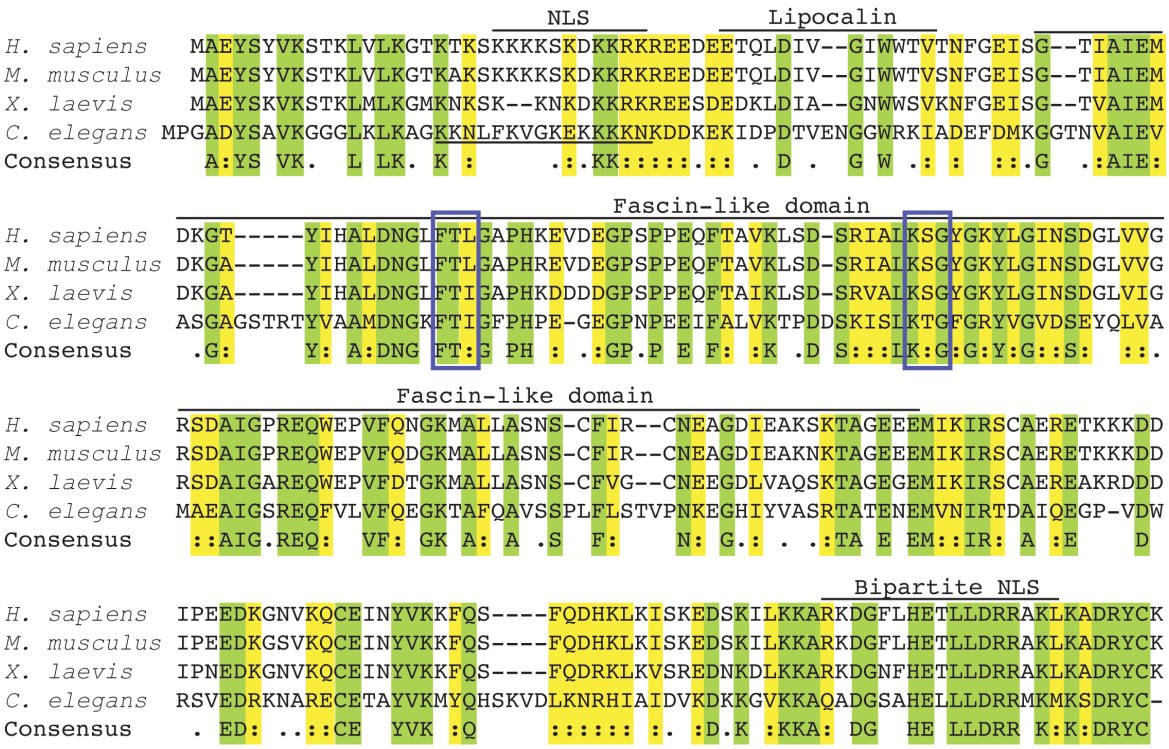

Figure A.4: FRG1 contains two highly conserved PKC sites. *In silico* sequence analysis revealed two potential phosphorylation sites (boxed region) in the fascin-like domain and they are highly conserved.

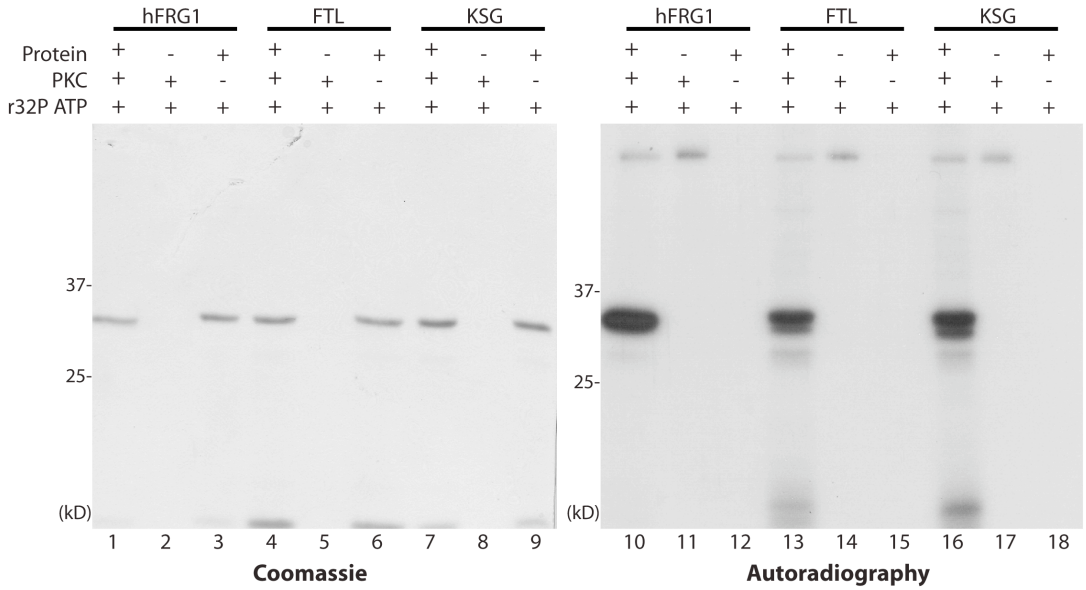

Figure A.5: Human FRG1 protein can be phosphorylated *in vitro*. Full-length FRG1 can be phosphorylated *in vitro* by Protein Kinase C (PKC). Mutating either FTL or KSG site leads to decreased phosphorylation level. Phosphorylation levels were visualized by autoradiography; proteins of input were visualized through Coomassie-Blue-staining of the same SDS-PAGE gel used for autoradiography.

Table A.1 PCR primers
Primers used for making fascin domain deletions and point mutations for human FRG1.

#1: CCAACATATGGCCGAGTACTCCTACGTGAAGTC

#2: CCAACTCGAGCTTGCAGTATCTGTCGGCTTTC

#3 (Δ58-71): ACCATTGTCGAGTGCGGCGGCCGCTGAAATTTCACCAAA

#4 (Δ58-71): TTTGGTGAAATTTCAGCGGCCGCCGCACTCGACAATGGT

#5 (Δ72-86): AGGGCCCTCATCAACGGCGGCCGCATGTATATAGGTTCC

#6 (Δ72-86): GGAACCTATATACATGCGGCCGCCGTTGATGAGGGCCCT

#7 (Δ87-102): GATTCTGGAATCAGAGGCGGCCGCTTCTTTGTGTGGAGC

#8 (Δ87-102): GCTCCACACAAAGAAGCGGCCGCCTCTGATTCCAGAATC

#9 (Δ103-118): TCCATCTGAATTTATGGCGGCCGCTAATTTGACAGCCGT

#10 (Δ103-118): ACGGCTGTCAAATTAGCGGCCGCCATAAATTCAGATGGA

#11 (Δ119-133): CCATTGTTCTCTTGGGGCGGCCGCACCAAGATATTTTCC

#12 (Δ119-133): GGAAAATATCTTGGTGCGGCCGCCCCAAGAGAACAATGG

#13 (Δ134-148): ATTTGAGGCCAACAAGGCGGCCGCTCCAATTGCATCTGA

#14 (Δ134-148): TCAGATGCAATTGGAGCGGCCGCCTTGTTGGCCTCAAAT

#15 (Δ149-162): TGCTTCTATGTCCCCGGCGGCCGCAGCCATTTTCCCATT

#16 (Δ149-162): AATGGGAAAATGGCTGCGGCCGCCGGGGACATAGAAGCA

#17 (Δ163-177): GGATCTAATCTTGATGGCGGCCGCTGCTTCATTGCATCT

#18 (Δ163-177): AGATGCAATGAAGCAGCGGCCGCCATCAAGATTAGATCC

#19 (FTL): TTTGTGTGGAGCTCCGGCGGCCGCAAGACCATTGTCGAG

#20 (FTL): CTCGACAATGGTCTTGCGGCCGCCGGAGCTCCACACAAA

#21 (KSG): AAGATATTTTCCATAGGCGGCCGCCAGGGCGATTCTGGA

#22 (KSG): TCCAGAATCGCCCTGGCGGCCGCCTATGGAAAATATCTT